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

Source	**Telecom ParisTech, Canon Research Centre France**
Status	**For consideration at the 106th MPEG meeting**
Title	**Storage of Tiled HEVC Video**
Author	Jean Le Feuvre, Cyril Concolato, Franck Denoual, Frédéric Mazé, Eric Nassor

1   Introduction

During the 105th MPEG Meeting, a framework for description of tiled HEVC bitstreams in ISOBMFF was proposed as a WD (cf W13670). At the same time, W13668 TuC has been issued on file format technologies for ordered combination of separated tracks.

This contribution reviews how these tools can be combined together for separate storage of HEVC tracks in order to provide easier extraction and addressing of spatio-temporal parts of the media.

2   Discussion

In the following sections, we describe different configurations of tiles, using a constant 2x2 tiling grid for the sake of clarity.


2.1   Independent Tiles

In this example, all tiles in the video bitstreams are coded with one or more slice segments per tile (i.e. no NAL unit covers two consecutive tiles). The tiles are independently decodable, as described in W13670.

2.1.1   Single track storage

When stored in a single track, the following structures will be found:
- Track with hvc1 sample description in stsd, containing all parameter sets
- 4 TileRegionGroupEntry entries in a sample group description (sgdp) of grouping_type 'trif', with groupIDs 1, 2, 3 and 4 (example numbers)
- 'nalm' sample group description with one entry, indicating that the first group of NALs (1 to L-1) maps to groupID 1, the second group of NALs (L to M-1) to groupID 2 and so

on. This sample group description box would typically have a default entry set to the first (and only) entry in the box, to indicate that all samples have the same subsample structure.
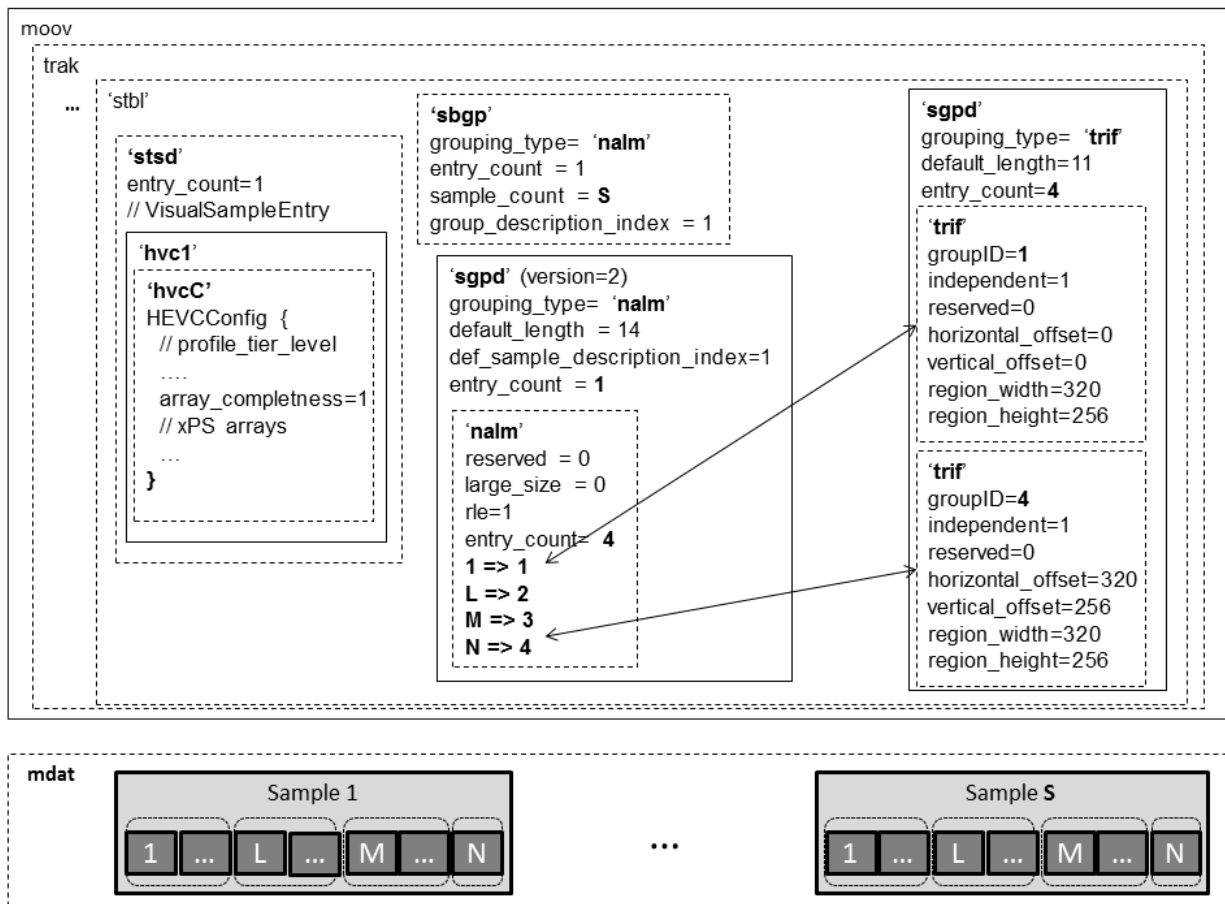
Optionally, there could be four TileSubTrackDefinition boxes if desired, to describe HEVC tile sub tracks.

Note: In this example, there are no NON-VCL NALUs in the sample, but it could be extended to take into account SEI messages for example.

2.1.2 One track per tile storage

When each tile is stored in a dedicated track, the following structures will be found:

- One track (trackID=10) with hvc1 sample description, with an hvcC containing all parameter sets. The track will typically have no samples if non-VCL units are not used, but this is not always the case.
- For each tile, one track (trackID=tile number) with:
  - hvt1 sample description, containing no hvcC. 'hvt1' code point indicate an HEVC tile track (definition will be given in next section).
  - 1 'trif' sample group description, with groupID equal to the tile number (example number). This sample group description would typically have a default entry set

to the first (and only) entry in the box, to indicate that all samples belong to the indicated tile.
  - o A 'tref' box with a reference of type 'sbas' to track 10, to indicate that the root of the dependency tree is on track 10
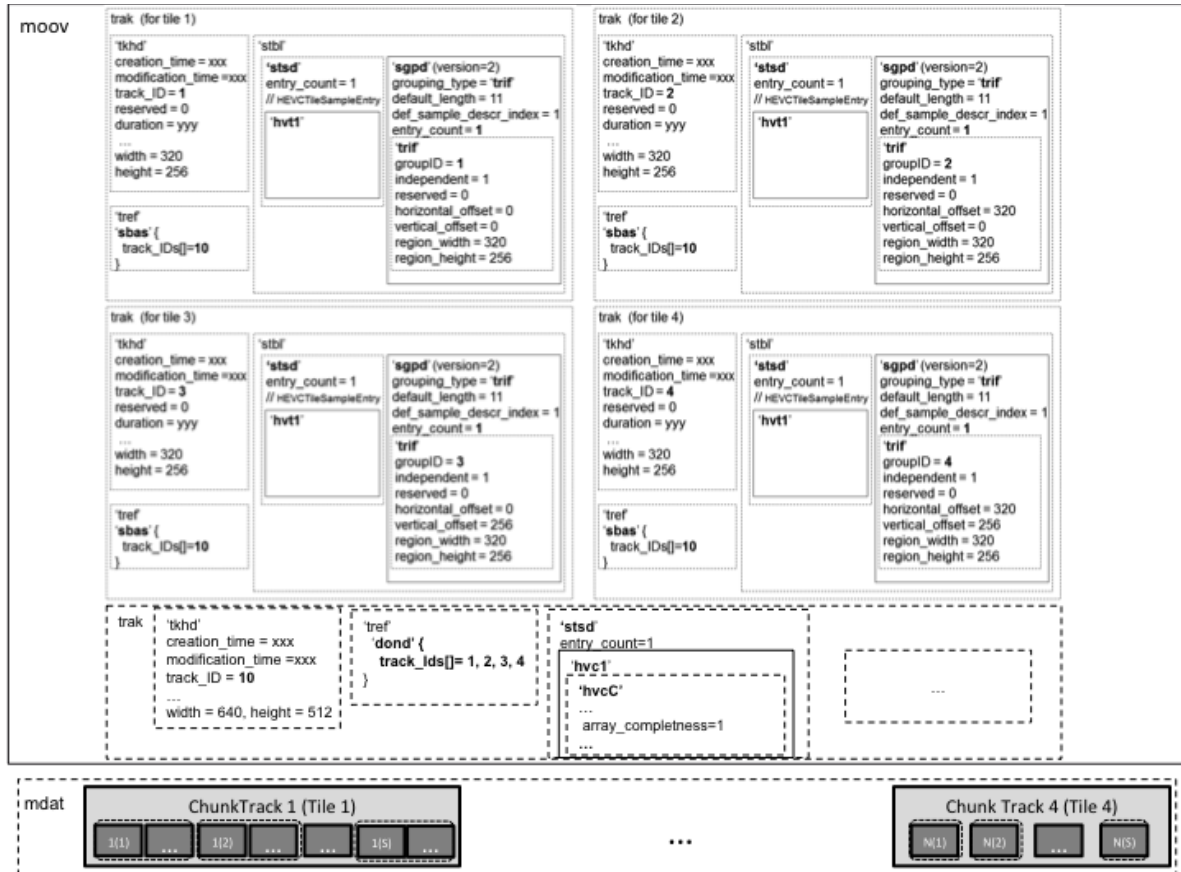
Furthermore, in order to indicate the original tile order, the hvc1 track will have a 'dond' track reference box (cf W13668). Note on the above figure, a possible tile-based data arrangement according to easily addressable tiles.
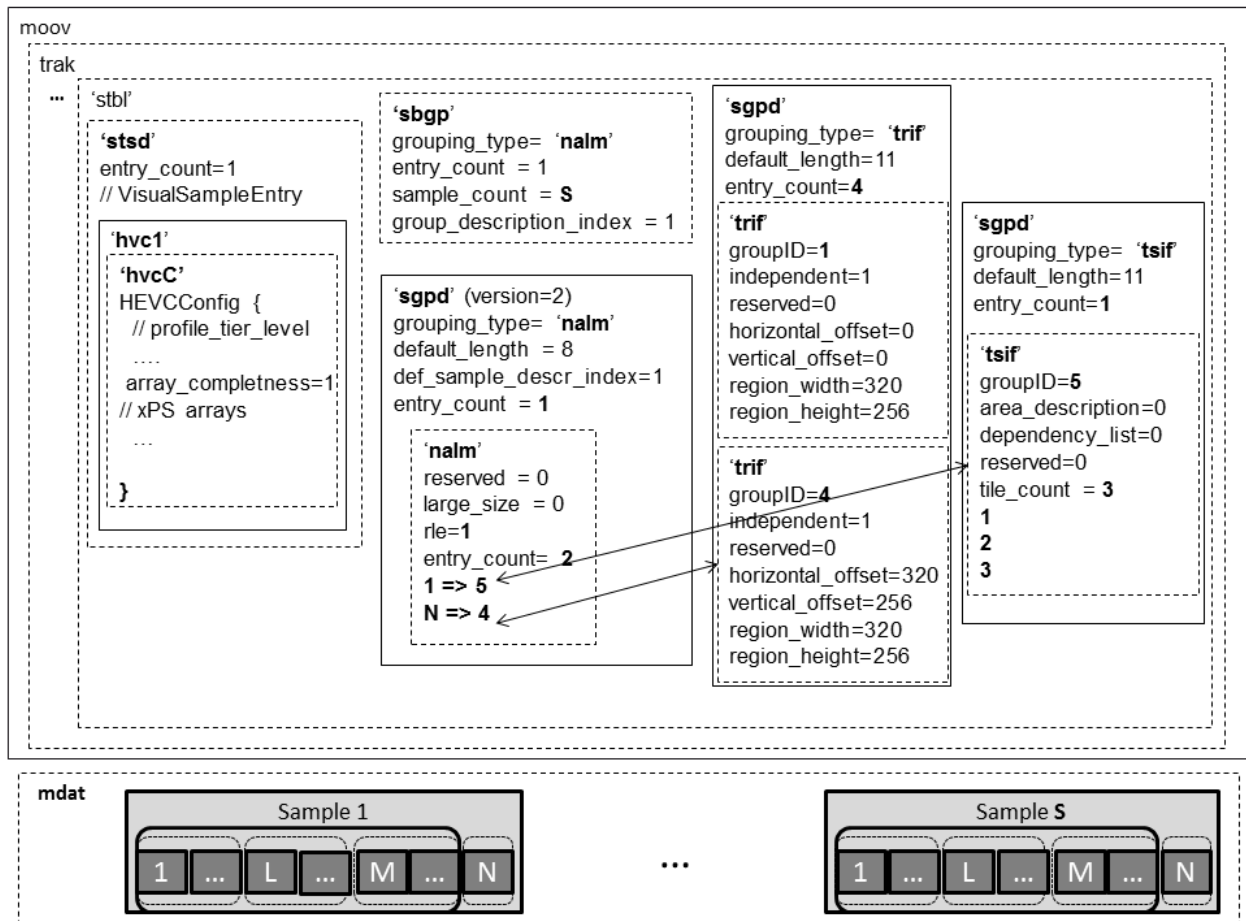
2.2 Dependent Tiles

In this example, we assume only tile 4 is independently decodable, and tiles 1 to 3 form an independently decodable TileSetRegion.

2.2.1 Single track storage

When stored in a single track, the following structures will be found:
  - Track with hvc1 sample description, containing all parameter sets
  - 4 'trif' sample group descriptions, with groupIDs 1, 2, 3 and 4 (example numbers)
  - 1 'tsif' sample group description with groupID 5, listing tileIDs 1, 2 and 3 in its dependency list
  - 'nalm' sample group description with one entry, indicating that the N-1 first NALUs maps to groupID 5, and N to groupID 4. This sample group description would typically have a default entry set to the first (and only) entry in the box, to indicate that all samples have the same subsample structure.

Optionally, there could be two TileSubTrackDefinition boxes if desired, to describe HEVC tile subtracks.

2.2.2 One track per independent tile/tileSet storage

When each independent tile/tileset is stored in a dedicated track, the following structures will be found:

- One track (trackID=10) with hvc1 sample description, with an hvcC containing all parameter sets. The track will typically have no samples if non-VCL units are not used, but this is not always the case.
- One track (trackID=1) with
    - o hvt1 sample description, containing no hvcC. 'hvt1' code point indicates an HEVC tile track (definition will be given in next section).
    - o 3 'trif' sample group description, with groupID equal to the tile number (example number).
    - o 1 'tsif' sample group description, with groupID to 5 and listing tiles 1, 2, and 3 in its dependency list. This sample group description would typically have a default entry set to the first (and only) entry in the box, to indicate that all samples belong to the indicated tileSet.
- One track (trackID=2) with
    - o hvt1 sample description, containing no hvcC. 'hvt1' code point indicate an HEVC tile track (definition will be given in next section).

o 1 'trif' sample group description, with groupID = 4. This sample group description would typically have a default entry set to the first (and only) entry in the box, to indicate that all samples belong to the indicated tile.
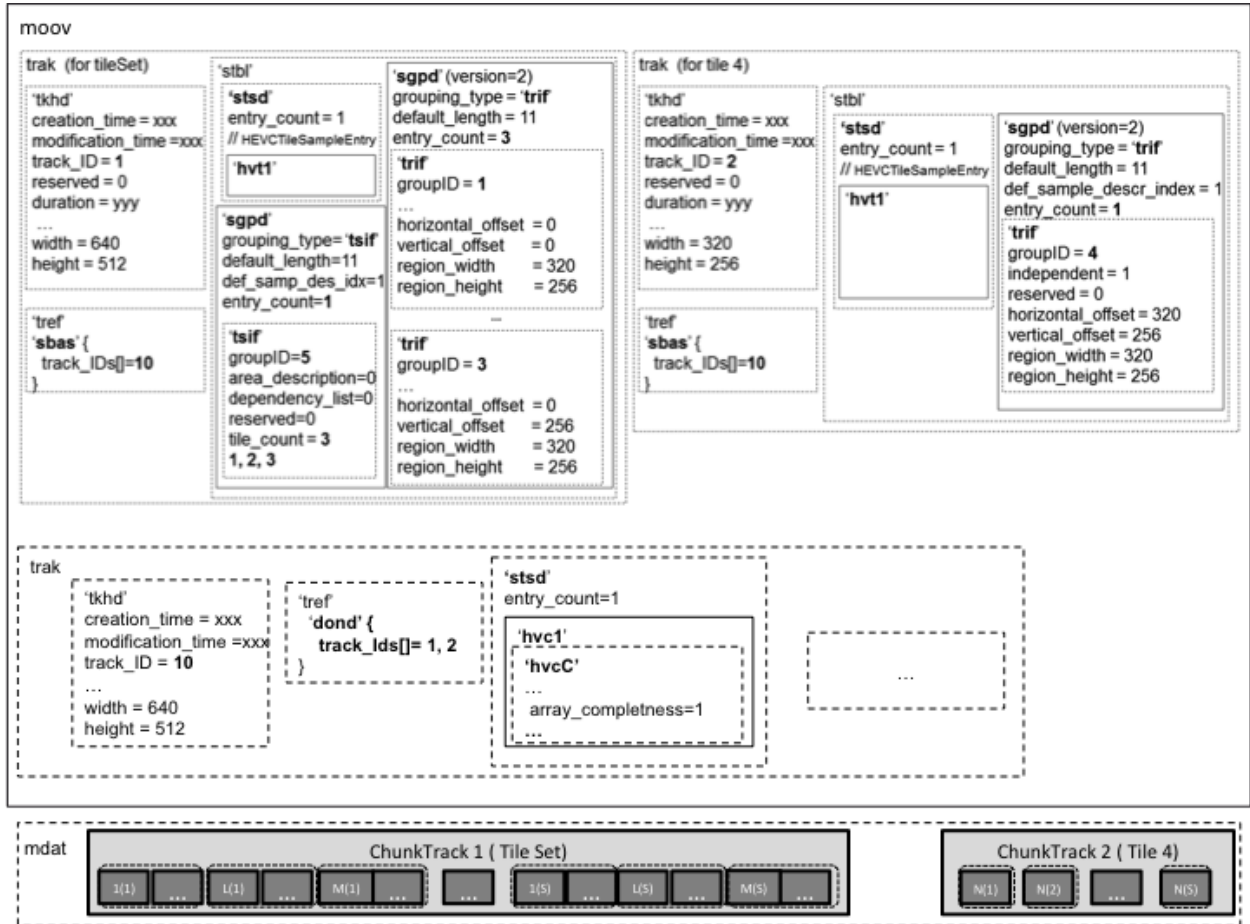
Furthermore, in order to indicate the original tile order, the hvc1 track will have a 'dond' track reference box (cf. W13668). The above figure provides a possible data arrangement to efficiently address tileSet or tile data.

3 Proposal

In order to allow for storage of HEVC tiles in dedicated tracks, we propose to amend the HEVC file format as follows:

Add new section

"8.4.10 HEVC Tiling"
There are cases where storing independently decodable HEVC tiles in different tracks may be useful for fast spatial and temporal access to the video content. For such cases, tracks may be created using HEVCTileSampleEntry sample description format.

A HEVC tile track is a video track for which there is either a 'dond' track reference from a base HEVC layer or a 'sbas' reference to the HEVC layer. The sample description type shall be 'hvt1'. Neither the samples in the tile track or the sample description box shall contain VPS, SPS or PPS NAL units, these NAL units shall be in the samples or in the sample description box of the track containing the base layer, as identified by the track references.

It is recommended to use 'dond' track references since they provide order information, which can be used to reconstruct the original bitstream without parsing slice headers to get the tile order.

A sample stored in a tile track is a complete set of slices for one or more tiles, as defined in ISO/IEC 23008-2. Typically if the track consists of a single tile, only the slice(s) used to code this tile will be found in the sample

An HEVC sample stored in a tile track is considered as a sync sample if the VCL NAL units in the sample indicate that the coded slices contained in the sample are Instantaneous Decoding Refresh (IDR) slices, Clean Random Access (CRA) slices, or Broken Link Access (BLA) slices.

Sub-sample and sample grouping defined for regular HEVC samples have the same definitions for an HEVC tile sample.
Note: When tiles of an HEVC video are stored in different tracks, there can be cases where no samples exist in the base layer. How and whether the tile samples are re-assembled to form a conformant HEVC bitstream is left up to the implementation.

The width and height of VisualSampleEntry for a tile track (sample entry type hvt1) shall be set to the maximum width and height of the Tile or TileSet contained in the track.

```
class HEVCTileSampleEntry() extends VisualSampleEntry ('hvt1'){

   extra_boxes          boxes;                    // optional
}
```

"

4   Conclusion

We propose including the definition of HEVC tile track in ISO/IEC 14496-15