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/M32284
January 2014, San Jose, USA

Source **Telecom ParisTech, Canon Research Centre France**
Status **For consideration at the 107th MPEG Meeting**
Title **Clarifications on HEVC Tile tracks**
Author Jean Le Feuvre, Cyril Concolato, Franck Denoual, Frédéric Mazé, Eric Nassor

1 Introduction

The WD for multi-layer HEVC and tiling (w13960) introduces the notion of Tile Tracks to store HEVC tiles in different tracks. The WD proposes to use extractors for the purpose of reconstructing the complete bitstreams, however the text needs further clarification for a better understanding of how tile tracks work. This contribution provides these clarifications.

2 Tile Track Definition

2.1 Track Dependencies

The current text in section 8.5.6.1 indicates:
"A HEVC tile track is a video track for which there is a 'sbas' reference to the HEVC layer"

In order to clarify the text for multi-layer extensions, it would be better to write:
"An HEVC tile track is a video track for which there is an 'sbas' reference to the HEVC track carrying the HEVC layer to which the tile(s) in this track belong. "

Furthermore, 'sbas' is used in SVC to indicate the base layer (AVC), and we can safely assume 'sbas' will be used in the same way for multi-layer HEVC. If we want to have a tile track design compatible with multi layer, we need to indicate the track carrying the associated HEVC layer, not the base layer. To avoid confusion when mixing tile storage and multi-layer storage, we suggest introducing a dedicated dependency type, 'tbas' for "tile base layer". The above sentence should then be reworded as follows:

"An HEVC tile track is a video track for which there is a 'tbas' reference to the HEVC track carrying NALUs of the HEVC layer to which the tile(s) in this track belong."

2.2 Tile Description

The text does not mention how the tile(s) is(are) described in the tile track. The tiling information is derived from the sample group information. We therefore suggest inserting the following sentence:

“A tile track shall either have:

- One and only one TileRegionGroupEntry and no TileSetGroupEntry
- One and only one TileSetGroupEntry and one or more dependent TileRegionGroupEntry this tileset is made of.

It is recommended to use default sample group mechanism to associate all the samples of the tile track to the appropriated TileRegionGroupEntry / TileSetGroupEntry.”

2.3 Tile Track SampleDescription

2.3.1 Visual Information

It is not clear from the current text how the various visual information other than width/height of the sample description in ‘*stsd*’ shall be treated. In particular:

- Presence of clean aperture box (‘*clap*’) or pixel sample aspect ratio box (‘*pasp*’)
- Width/height and layout information of the track header.

The decoding of a single tile track does not involve any layout operation, and the tile is decoded at the same place in the video decoder memory as if all tiles were decoded. We therefore suggest forcing the same track header information and no ‘*pasp*’/’*clap*’ in the sample description.

“The layout information in the track header (i.e., `layer` , `matrix` , `width` and `height`) of a tile track shall be identical to the track header information of the associated base track as identified by the ‘tbas’ track reference, and otherwise ignored. Any ‘*clap*’ and ‘*pasp*’ box in an ‘hvt1’ sample description shall be ignored.”

2.3.2 Bitrate

The current text in section 8.5.6.2.2 forbids any MPEG-4 Bitrate box in the ‘hvt1’ sample entry. It could however be useful to store per-tile bitrate in this box. We therefore suggest modifying the box as follows:

```
class HEVCTileSampleEntry() extends VisualSampleEntry (‘hvt1’){
   MPEG4BitRateBox ();              // optional
   extra_boxes          boxes;                  // optional
}
```

and modify the semantics of section 8.5.6.2.3 accordingly:

“The HEVCTileSampleEntry shall not contain any HEVCConfigurationBox or MPEG4ExtensionDescriptorsBox; these boxes are found in the base HEVC track sample description, as indicated by the ‘tbas’ track reference. Other optional extra boxes may be included. There shall be as many entries in the SampleDescriptionBox of an HEVC tile track as there are entries in the SampleDescriptionBox of the base HEVC track.”

2.4 Sample Format

The current text in section 8.5.6.1 states:
“An HEVC sample stored in a tile track is a complete set of slices for one or more tiles, as defined in ISO/IEC 23008-2.”

We believe it should be made more explicit that only slices can be in tile samples, which excludes parameter sets, SEI messages and other non-VCL nal units:
"An HEVC sample stored in a tile track is a complete set of slices for one or more tiles, as defined in ISO/IEC 23008-2, i.e. only the VCL NAL units forming those slices. This implies that the tile sample shall not contain any parameter set, SEI message or other non-VCL NAL units."

2.5 Extractors

Currently the WD does not forbid using extractors in the HEVC tile track to extract data from the base track; one potential use case could be to build a complete decodable bitstream for a single tile, including parameter sets and SEI messages from the base track. This approach has several drawbacks:

- The same extractors from $Tile_i$ to the base will be found in every tile sample; in the case where SEI prefix and suffix messages are present in the base track, this implies at least 2*12B = 4.8 kbps per tile track at 25Hz (this does not take into account any parameter set extraction in case of in-band storage)
- There could be empty samples in the base track which may be unsafe for existing implementations
- There is a need for either extractors from base to each tile track, or for an additional track using only extractors, to enable playback of all tile tracks together; in both cases, we end up with an additional authoring complexity (extra track or "crossed" extractors).

The situation is even more complex when tiled scalability is used, as will be shown in contribution m32285. We therefore suggest restricting the use of extractors only in non-tile tracks :

Replace in section 8.5.6.1

"The HEVC track may use extractors, as defined in annex B, to indicate how the original bitstream is reconstructed."

with

"An HEVC Tile track shall not use extractors. The track containing the associated layer, as indicated by the 'tbas' track reference, may use extractors, as defined in Annex B, to indicate how the original bitstream is reconstructed; in this case, this track will have a 'scal' track reference to each tile track from which it extracts data".

We also suggest adding a clarification on partial playback:

"Note: an implementation may decide to decode only a subset of the complete tiles of an HEVC sequence. In this case, it may use the tile dependency information in the TileRegionGroupEntry and TileSetGroupEntry sample group descriptions to discard un-needed tracks or ignore some extractors while decoding the HEVC track."

3 Conclusion

We suggest adding the proposed clarifications and modifications to the WD on multi-layer HEVC storage “Enhanced support of HEVC and MVC+D”.