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

ISO/IEC JTC1/SC29/WG11 MPEG2016/M39253
October 2016, Chengdu, CN

Source **Telecom ParisTech, Canon Research Centre France**
Status **For consideration during MPEG #116**
Title **Proposed Layered HEVC File Format Conformance Files**
Author Jean Le Feuvre, Cyril Concolato, Frederic Maze, Franck Denoual

1 Introduction

We have implemented some support for the Layered HEVC File Format in the GPAC Open Source Software (http://gpac.io), covering both tiling aspects and Layered aspects. This contribution proposes a set of files for conformance for the Layered HEVC File Format, and discusses some unclear points in the standard encountered while producing the files.

2 Discussion

2.1 Parameter Sets in hvc2 tracks

While the specification gives great details on where to find the parameter sets for a given sample, IRAP or not, there are cases where storage of parameter set is not optimal at all.

First note: the section 9.5.3.1.1 does not say anything about constraints on ‘hvc2’, only ‘lhv1’, ‘lhe1’ and ‘hev2’ are covered.

Consider a base+enhancement bitstream where the base layer is carried in its own ‘hvc1’ track and the enhancement layer is stored in a second track, parameter sets NOT in-band. If one decides to use extractors, two possibilities are offered by the spec:

- C1 (2 tracks): define an ‘hvc2’ track, where samples contain
  - extractors extracting from the base 'hvc1' track
  - enhancement layers NALUs
- C2 (3 tracks): define an ‘lhv1’ track containing the enhancement layer, and an ‘hvc2’ track where samples contain:
  - extractors extracting from the base ‘hvc1’
  - extractors extracting from the enhancement ‘lhv1’

In both cases, the ‘hvc2’ track shall contain both an ‘hvcC’ and an ‘lhcC’, but it is not clear from the specification if both of these config records shall contain the parameter sets of the layers (natively or not) present in the track or not. Several options exist for the config records of the ‘hvc2’ track:

- O1: forcing their presence: not very efficient due to the duplication of parameter sets

- O2: forbidding parameter sets: not desirable for use cases such as 'hvc2' tracks reconstructing bitstreams from a set of tiles, since the 'hvc2'/'hev2' tracks for these use cases may require their own SPS/PPS.
- O3: Check parameter set of the sample entry in the 'hvc2' track, and if not present, of the sample entry of the reference layer (similar as spec text for 'hev2'): it works well for C1, not for C2 since 'lhv1' track will not contain the reference layer but the extracted layer
- O4: Check parameter set of the sample entry in the 'hvc2' track, and if not present, of the sample entry of the extracted sample: it works well for both C1 and C2

We suggest going for option O4 (note that the same issue applies for temporal sublayers track, the base temporal sublayer cannot really be labeled "reference layer").

2.2 Parameter Sets in hev2 tracks

The spec currently says (for IRAP samples):
"For an IRAP picture of a given sample, track and layer, each parameter set needed for decoding the IRAP picture shall be included in one of the following:

a. the sample entry that applies to the given sample in the given track
b. the sample entry of the initial sample of a track carrying a reference layer of the given layer, where the initial sample is either the given sample's temporally collocated sample, when the temporally collocated sample contains an IRAP picture of the reference layer, or the previous sample that contains an IRAP picture of the reference layer

"

This is problematic since it does not give a mandatory order. In the case of 'hev2' track reconstructing an HEVC bitstream from a subset of the tiles of the base layer ('hev1' track), parameters sets in the 'hev2' sample description may be present and different from the parameter sets from the "reference layer" (see issue 2.2.2). Not mandating an order may result in players picking the parameter sets from the base 'hev1' rather than from the 'hev2', resulting in broken bitstream. It is furthermore not reasonable for a player to understand which one should be used, since these parameter sets may have the same IDs: it would require deep parsing of the slice header to understand which xPS should be used.

We would suggest mandating that 'hev2' and 'lhe1' have only in-band parameter sets (ie no xPS in the config record).

If this is not acceptable and if there are use cases for mixing both in-band and out-of-band parameter sets in 'hevc2', the specification should be fixed, to follow the following logic:

- if xPS is present in a sample, only use this xPS
- If no xPS is present in the sample, use the ones from this sample's sample description
- If no xPS in either sample or sample description,
    - If extractors are used, use the one from sample description of the extracted sample if no xPS is present in the extracted NALUs
    - Otherwise, use the one from sample description of the collocated sample of interest if no xPS is present in the collocated sample of interest

2.3 Sample Description Width/Height in hvc2/hev2

2.3.1 Multiple Layers

When an 'hvc2' track combines a base and a spatial enhancement layer, the base being extracted from an 'hvc1' track, the width and height in the sample description are set to the width and height of the base layer. This is:

- redundant with the width/height info already given in the base layer
- quite strange since it implies that we cannot build a file with a simple extractor setup describing the two resolutions

The only way to provide the information on both resolutions would be to add an extra 'lhv1' track and have the 'hvc2' track extracts from both, which is quite heavy… This might be problematic when generating DASH MPDs.

We suggest relaxing the constraint on width and height for 'hvc2'/'hev2' by saying that (in 9.5.3.1.2):
"If the sample entry is of type 'hvc1', 'hev1', 'hvc2', or 'hev2' and no LHEVCConfigurationBox is present in the sample description, the width and height documented in the VisualSampleEntry shall be set according to clause 4.6 using only the base layer information;"

2.3.2 Tile extracting tracks and base layer

For cases where an 'hvc2' track reconstructs a full HEVC bitstream from a subset of tiles, it is not clear what "base layer information" means in 9.5.3.1.2. In this case the base layer might be either the original HEVC bitstream (width / height indicated in sample description of track referenced through 'sbas'), or the newly constructed bitstream after resolution of extractors.
We believe the specification should be clearer here.

2.4 Presence of linf

The specification says "Every L-HEVC track, including the base track (when coded with HEVC), shall carry a 'linf' sample group".

This recommendation is too strong for the following cases

- all layers are in the same track, single track in the file: 'linf' should be optional in that case
- one base layer in one track and one (or more) enhancement layer(s) in a single other track: 'linf' should also be optional in that case: it is obvious that Layer-0 NALUs are in the base track and all other VCL Layer-N NALUs in the other track

We suggest modifying the specification as follows:

"Every L-HEVC track, including the base track (when coded with HEVC), shall carry a 'linf' sample group if there are more than one track with an 'sbas' track reference to the base layer track."

3 Conformance Files Description

The following bitstreams are provided, available for download at We provide such bitstreams at the location http://download.tsi.telecom-paristech.fr/gpac/MPEG/LHEVCFF/Conformance/. Note that these bitstreams may need changes depending on the resolutions above. For the current time, width/height is set accordingly to the spec and parameter sets are all cloned.

3.1 HEVC Conformance Files

- hvc1_only.mp4
  - Non-scalable, 'hvc1' single track (HEVC FF conformance)
- hev1_only.mp4
  - Non-scalable, 'hev1' single track (HEVC FF conformance)

3.2 HEVC Tiling Conformance Files

- hevc_tiles_single_track_nalm.mp4
  - independent 3x3 tiled HEVC single layer
  - one 'hvc2' track
  - "trif" and "nalm" with groupID=0 sample group descriptions
  - sample to group box with grouping_type="nalm" and grouping_type_parameter="trif"
- hevc_tiles_single_track_nalm_rle.mp4
  - independent 3x3 tiled HEVC single layer
  - one 'hvc2' track
  - sample to group box with grouping_type="nalm" and grouping_type_parameter="trif"
  - "trif" and "nalm" run-length encoded with groupID=0 sample group descriptions

- hevc_tiles_multiple_tracks.mp4
  - independent 3x3 tiled HEVC single layer
  - one 'hev1' and 9 'hvt1' tracks using "trif" with default sample group description
  - 'tbas' / 'sabt' track references for implicit reconstruction.
- hevc_tiles_single_track_trif_full_picture.mp4
  - single 'hvc1' track with 'trif' sample group description using full_picture flag
- hevc_tiles_single_track_nalm_all_intra.mp4
  - independent 3x3 tiled HEVC single layer
  - one 'hvc2' track
  - "trif" with independent_idc=2 and "nalm" with groupID=0 sample group descriptions
  - sample to group box with grouping_type="nalm" and grouping_type_parameter="trif"

3.3 L-HEVC SHVC Conformance Files

- shvc_hvc1_single_track.mp4
  - Simple LHEVC bitstream in 'hvc1' track
  - HEVC and LHEVC config
  - 'oinf' and 'linf' sample groups

- shvc_hev1_single_track.mp4
    - Simple LHEVC bitstream in ‘hev1’ track
    - all parameter sets are stored in-band
    - HEVC and LHEVC config
    - ‘oinf’ and ‘linf’ sample groups
- shvc_hvc2_single_track.mp4
    - Simple LHEVC bitstream in ‘hvc2’ track
    - HEVC and LHEVC config
    - ‘oinf’ and ‘linf’ sample groups
- shvc_hev2_single_track.mp4
    - Simple LHEVC bitstream in ‘hev2’ track
    - all parameter sets are stored in band
    - HEVC and LHEVC config
    - ‘oinf’ and ‘linf’ sample groups
- shvc_hvc1_hvc2_multiple_tracks_extractors.mp4
    - LHEVC bitstream in ‘hvc1’ track (base layer) and ‘hvc2’ track (enhancement layer)
    - HEVC and LHEVC config
    - ‘oinf’ and ‘linf’ sample groups Extractors, ‘oref’, ‘sbas’ and ‘scal’ track references
    - ‘hvce’ brand
- shvc_hev1_hev2_multiple_tracks_extractors.mp4
    - LHEVC bitstream in ‘hev1’ track (base layer) and ‘hev2’ track (enhancement layer)
    - HEVC and LHEVC config
    - ‘oinf’ and ‘linf’ sample groups
    - Extractors, ‘oref’, ‘sbas’ and ‘scal’ track references
    - all parameter sets are stored in-band
    - ‘hvce’ brand
- shvc_hvc1_lhv1_multiple_tracks_implicit.mp4
    - LHEVC bitstream in ‘hvc1’ track (base layer) and ‘lhv1’ track (enhancement layer)
    - LHEVC config in ‘lhv1’ only
    - ‘oinf’ and ‘linf’ sample groups
    - No extractors, ‘oref’ and ‘sbas’ track references
    - ‘hvci’ brand
- shvc_hev1_lhe1_multiple_tracks_implicit.mp4
    - LHEVC bitstream in ‘hev1’ track (base layer) and ‘lhe1’ track (enhancement layer)
    - LHEVC config in ‘lhe1’ only
    - ‘oinf’ and ‘linf’ sample groups No extractors, ‘oref’ and ‘sbas’ track references
    - all parameter sets are stored in-band
    - ‘hvci’ brand

3.4 Plan for L-HEVC Tiled Conformance Files

Conformance bitstreams for tiled layered HEVC will be produced for the next meeting due to some delay in SHM8+ conformant sequence generation. We plan to provide:

- non-tiled base hvc1/hev1 track with tiled hvc2/hev2 track (with extractors), trif information in all tracks and tile dependency to the full picture tile region of the base track
- non-tiled base hvc1/hev1 track with tiled lhv1/lhe1 track (implicit reconstruction), trif information in all tracks and tile dependency to the full picture tile region of the base track
- tiled base hvc1/hev1 with tiled lhv1/lhe1 (implicit reconstruction), trif information and tile dependency (enhancement tile to base tile)

4 Recommendation

We recommend to clarify the proposed aspects by issuing a COR to the L-HEVC file format and to add the contributed bitstreams, after potential alignment with said COR text, to the conformance suite of ISO/IEC 14496-15.