|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  30th Meeting: Gwangju, KR, 20–26 Jan. 2018 | Document: JCTVC-AD1011 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Issues affecting the usage of HEVC reference software for experimental studies** | | |
| *Status:* | Output document approved by JCT-VC | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | **Gary Sullivan** Microsoft Corp. 1 Microsoft Way Redmond, WA 98052 USA  **Jens-Rainer Ohm** Institute of Communication Engineering RWTH Aachen University Melatener Straße 23 D-52074 Aachen | Tel: Email:  Tel: Email: | +1 425 703 5308 [garysull@microsoft.com](mailto:garysull@microsoft.com)  +49 241 80 27671 [ohm@ient.rwth-aachen.de](mailto:ohm@ient.rwth-aachen.de) |
| *Source:* | JCT-VC | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

Discussions were held at the 30th JCT-VC meeting (Gwangju, Korea, January 2018) to determine the extent to which there are issues affecting the usage of HEVC reference software for experimental purposes. The expressed concerns were discussed to determine the extent to which they are valid and to determine whether it is feasible to extend the software to address the limitations. The identified issues are listed, along with remarks recorded for consideration of each issue.

# List of specific issues and responses

Document JCTVC-AD0025 contains questions about HEVC reference software capabilities and limitations for experimental uses. It was generated in response to comments and queries that arose in MPEG (see WG 11 output documents N 16522 and N 17133, and inputs m42109 and m42150).

It was requested to determine the extent to which these concerns are valid and whether it is feasible to extend the software to address the limitations. The identified issues are listed below, along with remarks recorded for consideration of each issue.

1. 3D-HEVC (and perhaps MV-HEVC) encoding/decoding with more than 16 views (suggested to be a non-normative issue; see patch provided in WG 11 N 16522)

Remarks: This behaviour would not be consistent with original intent, see under 3. The decoder reportedly uses an assert, which just doesn't seem like it should be there, and the encoder reportedly doesn't really have a problem in this regard.

1. Encoding/decoding with more than 16 PPSs, referring to the constraint that "pps\_seq\_parameter\_set\_id shall be in the range 0 to 15 inclusive " (MV-HEVC and 3D-HEVC; it was suggested that sharing PPSs across multiple views may be an acceptable work-around)

Remarks: According to the name of the specific syntax element that is referred to here, this is actually about SPSs, not PPSs. Sharing SPSs is the normal practice, so it seems there may not really be an issue in this regard.

1. Maximum layer ID range "The value of nuh\_layer\_id shall be in the range of 0 to 62, inclusive." (MV-HEVC and 3D-HEVC). For experiments reported in MPEG input document m42150, this limitation was addressed by increasing the length of nuh\_layer\_id from 6 bits to 8 bits, and removing 2 bits from nuh\_temporal\_id\_plus1 (which became only one bit). The latter change was motivated by not wanting to increase the 16-bit length of the NAL unit header.

Remarks: We had already planned to use 63 as an escape code (and there is a NOTE in the spec about it). There was a prior proposal JCT3V-F0071 from HHI about how to handle this.

1. Lack of support for temporal sub-layers in the HTM software (3D-HEVC).

Remarks: This may be just an encoder issue. It was remarked that this feature is not necessarily important to have. It is unclear how much effort it would be to implement (it may be an encoder issue or more), but would not be important to have for lightfield array experiements.

1. Lack of support (or limited support) for vertically displaced inter-view prediction in 3D-HEVC.

Remarks: This is not a problem for MV-HEVC; only 3D-HEVC. The software will run, but is just suboptimal for vertically-displaced views. This affects normative content, as there is a conversion from depth to disparity and this has only been designed for horizontal disparity. This might also be solved by disabling tools that rely on horizontal disparity (VSP). Having this better supported would probably help coding efficiency, although it's not clear how much (probably not a large amount).

1. depthRefinementFlag limited to 0 for views other than those with view ID equal to 0, 1, or 2 in 3D-HEVC.

Remarks: This may affect the encoder and decoder. There is no corresponding restriction in the spec/design, so to some extent this may be a bug in terms of existing intended functionality. It does not seem especially important to fix.

1. Depth maps with a bit depth greater than 8 in 3D-HEVC.

Remarks: This is not a problem for MV-HEVC; it would be nice to have the functionality also for 3D-HEVC. 10 bit should be feasible (but not for all specific depth coding tools of 3D HEVC). Integration of the HTM with RExt would be a major effort

1. Views are required to be ordered strictly from left to right (or vice-versa) for HTM encoding for 3D-HEVC (view IDs must strictly increase (or strictly decrease) and the encoder needs this relationship for effective encoding).

Remarks: This is primarily an encoder problem of the HTM. There is no support for camera parameters that are vertically displaced. This has some relationship to item 5. It was said that this may not be so difficult to solve. Basically, it is already solved in the MPEG input document m16522.

The first important steps should be items 1 and 3 (increasing number of views).

Second would be item 7 (increasing bit depth).

Third would be vertical disparity (item 5).

1. Operation with picture sizes larger than 8192×4320 (see MPEG input document m42109).

Remarks: There may not really be a problem in the reference software in this regard. However, we should have some ability to use an unspecified profile/level for experimental purposes. This may not be in all codebases. It is likely solvable when changing software parameters before compiling. That may be the only problem.

1. Operation with bit depths greater than 12, and especially equal to 16, including lossless modes (see MPEG input document m42109).

Remarks: It seems that the reporters of this issue did not use the reference software. This is not a problem for the HM (especially if compiled for high bit depth operation, in which case the performance will be better), but it is a problem for SHM & HTM. Properly, for anything more than 12 bits, extended precision processing should be used, but extended precision processing is not supported in SHM & HTM.

Items 9 and 10 arrived as comments about experimental usage of a software implementation that is *not* our reference software. These comments are also not about multiview coding – just ordinary single-layer coding.

It was *agreed* that enabling experimental uses is desirable.

Fixing some items like 1 & 3 would enable significant experimentation without necessarily being difficult to enable.

We should identify which codebases are relevant. Candidates are the HM, SHM, HTM, and SCM codebases. HM and SCM don't have multiview support. We should enhance all of these that we can.

The tough items (where the functionality is both relatively desirable and difficult to enable) are items 4 and 8. Item 10 is also difficult if the high bit depth support is needed in combination with scalable or multiview capability.

Further work on this is to be conducted in JCT-VC AHG3.