|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  37th Meeting: Geneva, CH, 4–10 October 2019 | Document: JCTVC-AK0027 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **On HRD for splicing: Bug-fix** | | |
| *Status:* | Input document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Yago Sanchez  Robert Skupin  Karsten Suehring Thomas Schierl | Email: | {firstname.lastname}@hhi.fraunhofer.de |
| *Source:* | Fraunhofer HHI | | |

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

# Abstract

This contribution provides details of an asserted bug of the Hypothetical Reference Decoder (HRD) model of HEVC for spliced bitstreams. The contribution proposes a bug-fix for the reported issue.

# Problem Statement

During the last JCT-VC and JVET meeting, two contributions JCTVC-AJ0026 and JVET-O0496 were discussed pointing out the potential issue in Equation C-10 of the HEVC and VVC specification for bitstream splicing.

In those cases, when concatenationFlag is set to 1, that the AuNominalRemovalTime for the splicing point under some circumstances (non-seamless splicing) may be higher than   
AuNominalRemovalTime[ prevNonDiscardablePic ]+ ( auCpbRemovalDelayDeltaMinus1 + 1 ).

The text in HEVC specifies:

if( !concatenationFlag ) {  
 baseTime = AuNominalRemovalTime[ firstPicInPrevBuffPeriod ]  
 tmpCpbRemovalDelay = AuCpbRemovalDelayVal  
} else {  
 baseTime = AuNominalRemovalTime[ prevNonDiscardablePic ]  
 tmpCpbRemovalDelay =  
 Max( ( auCpbRemovalDelayDeltaMinus1 + 1 ), (C‑10)  
 Ceil( ( InitCpbRemovalDelay[ SchedSelIdx ] ÷ 90000 +  
 AuFinalArrivalTime[ n − 1 ] − AuNominalRemovalTime[ n − 1 ] ) ÷ ClockTick ) )  
}  
AuNominalRemovalTime[ n ] = baseTime + ClockTick \* ( tmpCpbRemovalDelay − CpbDelayOffset )

In the following figure, it is shown a splicing operation that does not allow for seamless splicing due to the fact that the InitCpbRemovalDelay[ SchedSelIdx ] imposes a later removal time after splicing than equidistant CPB removal times as present in seamless playout.



If instead of deriving the removal time of AU 6 from the AU n-1 (i.e. AU 5) as indicated above in the figure, we use the last non-discardable picture (AU 4) it would result in the following:

taf6= trm4 + >= trm6 (CPB underflow resulting in non-seamless splicing)

trm6= taf5+InitCpbRemovalDelay

=taf6 – trm4= Size/Bitrate – (trm4 – taf5)

trm4+Size/Bitrate – (trm4 – taf5)>= taf5+InitCpbRemovalDelay

Size/Bitrate – (trm4 – taf5)>= (taf5 – trm4)+InitCpbRemovalDelay

taf5+ Size/Bitrate = **taf6 >= trm4 + (taf5 – trm4)+InitCpbRemovalDelay**

As seen in the derived formula, the part marked in yellow in C-10 above is not correct. Therefore, a fix is proposed in this contribution.

The current formula in the specification was proposed in JCTVC-L0328, which was based on an earlier contribution for AVC JVT-V055. In the original contribution for AVC, the removal time of the splicing point was based on the removal of the directly preceding AU in decoding order as shown in the following equations:

tr,n( n ) = tr,n( n-1 ) + tc \* spliceDelay    (1)  
  
with spliceDelay being the maximum value between the 1-frame delay for a given framerate and a value based on the initial\_cpb\_removal\_delay of the splicing point. That value reflects that the removal time of the splicing point cannot be smaller than the time instant corresponding to initial\_cpb\_removal\_delay after the final arrival time of the directly preceding AU in decoding order. This needs to be ensured so that the HRD model for the bitstream after the splicing point is fulfilled.

spliceDelay = Max(DeltaTfiDivisor, Ceil( (initial\_cpb\_removal\_delay[ SchedSelIdx ] ÷ 90000 +  
 taf( ns - 1 ) – tr,n( ns - 1 )) ÷ tc )      (2)

The difference in JCTVC-L0328 and therefore in HEVC compared to JVT-V055 is that the removal delays are not computed based on the removal time of the directly preceding AU in decoding order but on the removal time of the previous nonDiscardable picture in order to make the design robust to the removal of pictures from higher temporal sub-layers.

This feature can be applied to the case where seamless splicing occurs. I.e., the removal time can be expressed as:

tr,n(n) = AuNominalRemovalTime[ prevNonDiscardablePic ] + tc \*( auCpbRemovalDelayDeltaMinus1 + 1 ) (3)

However, for non-seamless splicing it still needs to be guaranteed that the removal time of the splicing point cannot be smaller than the time instant corresponding to initial\_cpb\_removal\_delay after the final arrival time of the directly preceding AU in decoding order. And this is expressed as:

tr,n( n ) = AuNominalRemovalTime[ n-1 ] + tc \*Ceil( (initial\_cpb\_removal\_delay[ SchedSelIdx ] ÷ 90000 +  
 taf( ns - 1 ) – tr,n( ns - 1 )) ÷ tc )     (4)

As shown above, if the nonDiscardable picture is taken as a reference for the derivation of the removal time when non-seamless splicing happens, the derived removal time is wrong. The maximum value of the two above ( 3 and 4) is the one that needs to be derived form Equation C-10, which was the solution proposed in JCTVC-AJ0026 and JVET-O0496 and replicated below.

if( !concatenationFlag ) {  
 baseTime = AuNominalRemovalTime[ firstPicInPrevBuffPeriod ]  
 tmpCpbRemovalDelay = AuCpbRemovalDelayVal  
} else {  
 baseTime1 = AuNominalRemovalTime[ prevNonDiscardablePic ]  
 tmpCpbRemovalDelay1 = ( auCpbRemovalDelayDeltaMinus1 + 1 )  
 baseTime2 = AuNominalRemovalTime[ n − 1 ]  
 tmpCpbRemovalDelay2 =   
 Ceil( ( InitCpbRemovalDelay[ SchedSelIdx ] ÷ 90000 +  
 AuFinalArrivalTime[ n − 1 ] − AuNominalRemovalTime[ n − 1 ] ) ÷ ClockTick )   
 (C‑10)  
 if( baseTime1 + ClockTick \* tmpCpbRemovalDelay1 <   
 baseTime2 + ClockTick \* tmpCpbRemovalDelay2 ) {  
 baseTime = baseTime2  
 tmpCpbRemovalDelay = tmpCpbRemovalDelay2  
 } else {  
 baseTime = baseTime1  
 tmpCpbRemovalDelay = tmpCpbRemovalDelay1  
 }  
}  
AuNominalRemovalTime[ n ] = baseTime + ClockTick \* ( tmpCpbRemovalDelay − CpbDelayOffset )

# Patent rights declaration(s)

**Fraunhofer HHI may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**