|  |  |
| --- | --- |
| **Joint Collaborative Team on 3D Video Coding Extension Development**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  3rd Meeting: Geneva, CH, 17–23 Jan. 2013 | Document: JCT3V-C0062 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **AHG7: Parallel decoding SEI message for MV-HEVC** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Ying Chen Vadim Seregin Adarsh K. Ramasubramonian Li Zhang Ye-Kui Wang  5775 Morehouse Drive San Diego, CA 92121 USA | Tel: Email: | 1-858-845-6589 [cheny@qti.qualcomm.com](mailto:cheny@qti.qualcomm.com)  [vseregin@qti.qualcomm.com](mailto:vseregin@qti.qualcomm.com)  [aramasub@qti.qualcomm.com](mailto:aramasub@qti.qualcomm.com)  [lizhang@qti.qualcomm.com](mailto:lizhang@qti.qualcomm.com)  [yekuiw@qti.qualcomm.com](mailto:yekuiw@qti.qualcomm.com) |
| *Source:* | Qualcomm Incorporated | | |

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

# Abstract

A parallel decoding SEI message for MV-HEVC similar to the parallel decoding SEI message in MVC is proposed with three modifications: 1) the delay required for parallel decoding is signalled in unit of coding tree units (CTU); 2) the horizontal CTU delay is also signalled to avoid large increased delay caused by the size of a CTU; 3) the delay is signalled once for all the non-base views and is applicable to any view that utilizes inter-view prediction.

# Introduction

## Example of the LCU delay in horizontal and vertical directions

There is a delay of CTUs if it is desired to decode base view and enhancement views in parallel. In Fig.1, base and enhancement views are depicted. Assume that the current block in the enhancement view is using inter-view prediction with a disparity motion vector MV.

In the beginning, let’s assume that the in-loop filtering processes like deblocking filter and SAO are performed on CTU (also known as largest coding unit, LCU) basis.

To get the full reconstructed pixels in the base view to be used for inter-view prediction, the pixels in the referenced LCU should be fully reconstructed, meaning in-loop filtered. To achieve this, the immediate neighboring LCUs have to be at least de-blocked, and the neighboring LCUs for those de-blocked LCUs have to be at least inverse transformed (e.g., after motion compensation). So, for the case that decoder processes are performed on LCU basis there is a delay of three LCU rows in the vertical direction for the enhancement view decoding, unless the vertical component of MV is always 0. In other words, to start decoding the enhancement view at least three LCU rows of the base view should be available even if the MV has a very small magnitude. The LCU row containing the inter-view prediction has to have full reconstructed data, the next LCU row has to have at least de-blocked pixels, and the third LCU row has to have at least the pixels after inverse transform.

Let’s simplify it to the case that three above mentioned LCU rows require full reconstruction. Assume the LCU address of the reference block is (LCUaddrX, LCUaddrY). In this case, the part of the base view picture located, starting from the LCUaddrY + 2, cannot be used for inter-view prediction.

Doing the same analysis, assuming that there is a horizontal restriction for the inter-view prediction, the part of the base view picture located after LCUaddrX + 2 cannot be referred for inter-view prediction.

If reference block is located at the LCU boundary in the base view, then the neighboring LCUs have to have full reconstructed pixels that will be necessary for the interpolation.

So, the delay can be increased by one in horizontal and vertical directions, i.e. LCUaddrY + 3 and LCUaddrX + 3.



Fig. 1. Illustration of the LCU rows delay.

## Discussions on impact on decoder implementations

In above section, there is an assumption that decoding processes are performed on LCU basis. If we assume that the decoding process can be done line by line, then the unreferenced base view picture can be expressed in the number of lines and columns. The part of the base view that should not be referenced starts from the point (BottomRightX + 8, BottomRightY + 8) with the coordinates in horizontal and vertical directions, since 4 pixels are need for interpolation, 3 pixels for deblocking and 1 pixel for SAO. Here, BottomRight is the point at the bottom right corner of the referenced block in the base view. It is noticed that in this case, two LCU rows delay may be typically sufficient for parallel decoding if the magnitude of disparity motion vectors are relatively small.

# Proposal

It is proposed that the MVC parallel decoding SEI message is simplified and used for MV-HEVC, by allowing signaling of the horizontal delay. Similar to the parallel decoding SEI message in MVC, it is assumed such an SEI message is present and applicable when raster scan decoding order applies.

## Parallel decoding information SEI message syntax

|  |  |
| --- | --- |
| parallel\_decoding\_info( payloadSize ) { | Descriptor |
| **video\_parameter\_set\_id** | ue(v) |
| **pdi\_init\_delay\_ctu\_vertical\_minus2** | ue(v) |
| **pdi\_init\_delay\_ctu\_horizontal\_minus2** | ue(v) |
| } |  |

## Semantics

**video\_parameter\_set\_id** specifies a video parameter set that contains the inter-view dependency relationship information. The value of video\_parameter\_set\_id shall be equal to the value of video\_parameter\_set\_id referenced by a view component of the coded picture of the access unit containing the parallel decoding information SEI message.

**pdi\_init\_delay\_ctu\_vertical\_minus2** and **pdi\_init\_delay\_ctu\_horizontal\_minus2** specify the unavailable reference area in any reference view component that shall not be used for inter-view reference by the coded view component that uses inter-view prediction as specified in the active video parameter set identified by video\_parameter\_set\_id.

The variable horCtb is derived as follows.

horCtb = pdi\_init\_delay\_ctu\_horizontal\_minus2   
 ? pdi\_init\_delay\_ctu\_horizontal\_minus2 + 2 + ( CtbAddrInRS % PicWidthInCtbsY ) : 0

The variable verCtb is derived as follows.

verCtb = CtbAddrInRS / PicWidthInCtbsY + pdi\_init\_delay\_ctu\_vertical\_minus2 + 2

The variable refCtbAddr that denotes the address of refCtb is derived as follows.

refCtbAddr =Min( horCtb, PicWidthInCtbsY ) + verCtb \* PicWidthInCtbsY .

The unavailable reference area includes all the coding tree units with an address equal or larger than refCtbAddr. When decoding the coded view component, samples from the unavailable reference area from the view component of a reference view shall not be referred to by the inter-view prediction process.

# Discussions

Due to different decoder implementations, more precise signaling of the delay might not benefit all conforming decoders from the perspective of decoding delay for parallel decoding. So, it is safe to always assume the worst case. In other words, although more accurate signaling of delay might be helpful for certain decoders, it need not be specified in the mandatory part of the bitstream since it is difficult to be used by the normative decoding process.

# Patent rights declaration(s)

**Qualcomm Incorporated 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).**