|  |  |
| --- | --- |
| **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-C0149 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **3D-CE5.h related:** **Explicit** **signaling of the second** **collocated picture for 3D-HEVC** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Kai Zhang, Jicheng An, Shawmin Lei  MediaTek Beijing North Building 10F, Raycom Infotech Park Tower C, No. 2 Kexueyuan South Rd., Haidian District, Beijing, China 100190 | Tel: Email: | Shawmin Lei +886-3-5670766 ext. 25555 {kai.zhang, jicheng.an, shawmin.lei}@mediatek.com |
| *Source:* | MediaTek Inc. | | |

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

# Abstract

In the current HTM, a second collocated picture is used in the DV derivation process, where the second collocated picture is derived implicitly and dependent on the temporal ID in the NAL header. In order to unify with the first collocated picture and provide more encoder flexibility, it is proposed to signal the reference index of the second collocated picture in the slice header explicitly as the first collocated picture is also coded in the slice header. It is reported that the explicit signaling does not cause any coding efficiency loss for coded and synthesized videos.

# Introduction

Disparity Vector (DV) plays a very important role in the current HTM [1]. Besides being used in inter-view prediction directly, DV is also utilized in inter-view motion vector (MV) prediction and inter-view residual prediction. In the inter-view MV prediction and inter-view residual prediction, DV is estimated instead of being sent from the encoder to the decoder. In the current HTM, the estimated DV is derived from the DV in spatial and temporal neighboring blocks [2]. DV-MCP [3], which uses DVs used for inter-view MV prediction in neighboring merge blocks in the DV derivation process, has also been adopted into the current HTM.

To utilize temporal neighboring blocks sufficiently, HTM currently uses two collocated pictures in the DV derivation process [4]. Temporal neighboring blocks in the two collocated pictures are checked in an order to find the candidate DV in the DV derivation process. The first collocated picture is the same as the one used in the Temporal Motion Vector Prediction (TMVP) process. The second collocated picture is another reference picture, which is different from the first collocated picture. Unlike the first collocated picture, which is signaled explicitly in the slice header, the second collocated picture is derived implicitly both on the encoder and the decoder following the same rule.

The second collocated picture is derived in the reference picture lists with the ascending order of reference picture indices, and added into the candidate list, given as follows:

1. A random access point (RAP) is searched in the reference picture lists. If found, the RAP is placed into the candidate list for the second picture and the derivation process is completed. In a case that the RAP is unavailable for the current picture, go to step (2).
2. A picture with the lowest temporal ID (TID) is searched out and placed into the candidate list of the temporal pictures as the second entry.
3. If multiple pictures with the same lowest TID exist, a picture of less POC difference with the current picture is chosen.

# Proposed method

During the standardization of HEVC version 1, the first collocated picture was implicitly derived, and only one collocated picture is used for each slice in HEVC format. Although the implicit derivation worked properly in common test conditions, it is not guaranteed to work well in all other conditions. In order to provide more encoder flexibility, a straightforward method, which explicitly signals the reference index of the collocated picture in slice header, was tested and found no coding efficiency loss in comparison with the implicit method. Therefore, the latest HEVC standard uses the explicit signaling for the collocated picture.

Based on the above reason, for 3D-HEVC, we propose to explicitly signal the reference index of the second collocated picture in slice header as the first collocated picture is also coded in slice header.

# Experimental results

The proposed explicit signaling is implemented on HTM-5.0.1[5] and tested under the common test condition [6]. The results are shown in Table 1, where the impact on coding efficiency is negligible.

**Table 1. Results of the proposed explicit signaling for the second collocated picture**



# Working draft modifications

The proposed working draft modifications are as follows.

|  |  |
| --- | --- |
| slice\_header( ) { |  |
| … |  |
| if( slice\_temporal\_mvp\_enable\_flag ) { |  |
| if( slice\_type = = B ) |  |
| **collocated\_from\_l0\_flag** | u(1) |
| if( ( collocated\_from\_l0\_flag && num\_ref\_idx\_l0\_active\_minus1 > 0 )  | | ( !collocated\_from\_l0\_flag &&  num\_ref\_idx\_l1\_active\_minus1 > 0 ) ) |  |
| **collocated\_ref\_idx** | ue(v) |
| } |  |
| if(slice\_temporal\_mvp\_enable\_flag &&  (multi\_view\_mv\_pred\_flag || multi\_view\_mv\_pred\_flag)){ |  |
| if( slice\_type = = B ) |  |
| **second\_collocated\_from\_l0\_flag** | u(1) |
| if( ( collocated\_from\_l0\_flag && num\_ref\_idx\_l0\_active\_minus1 > 0 )  | | ( !collocated\_from\_l0\_flag &&  num\_ref\_idx\_l1\_active\_minus1 > 0 ) ) |  |
| **second\_collocated\_ref\_idx** | ue(v) |
| } |  |
| … |  |
| } |  |

**second\_****collocated\_from\_l0\_flag** equal to 1 specifies the second picture that contains the collocated partition in the derivation process for a disparity vector is derived from list 0, otherwise the picture is derived from list 1.When second\_collocated\_from\_l0\_flag is not present, it is inferred to be equal to collocated\_from\_l0\_flag.

**second\_collocated\_ref\_idx** specifies the reference index of the second picture that contains the collocated partition in the derivation process for a disparity vector. When the current slice is a P slice, second\_collocated\_ref\_idx refers to a picture in list 0. When the current slice is a B slice, second\_collocated\_ref\_idx refers to a picture in list 0 if second\_collocated\_from\_l0 is 1, otherwise it refers to a picture in list 1. second\_collocated\_ref\_idx shall always refer to a valid list entry, and the resulting picture shall be the same for all slices of a coded picture. When second\_collocated\_ref\_idx is not present, it is inferred to be equal to collocated\_ref\_idx.

Derivation process for the candidate picture list for disparity vector derivation

The variable NumDdvCandPics is set equal to 0 and the candidate picture list DdvCandPicList with a number of NumDdvCandPics elements is constructed as follows.

When slice\_temporal\_mvp\_enable\_flag is equal to 1 the following ordered steps apply:

1. DdvCandPicList[ 0 ] is set equal to RefPicListX[ collocated\_ref\_idx ], with X equal to 1-collocated\_from\_l0\_flag~~, and NumDdvCandPics is set equal to 1~~.
2. DdvCandPicList[ 1 ] is set equal to RefPicListX[ second\_collocated\_ref\_idx ], with X equal to 1-second\_collocated\_from\_l0\_flag. ~~The variable lowestTemporalIDRefs is set equal to 7.~~
3. If PicOrderCnt(DdvCandPicList[ 0 ]) equals to PicOrderCnt(DdvCandPicList[ 1 ]), NumDdvCandPics is set equal to 1. Otherwise, it is set equal to 2. ~~NumDdvCandPics, DdvCandPicList[ 1 ] and lowestTemporalIDRefs are derived as specified in the following.~~

~~for ( dir = 0; dir < 2 ; dir++) {  
 X = dir ? collocated\_from\_l0\_flag : (1- collocated\_from\_l0\_flag)  
 for ( i =0 ; i <= num\_ref\_idx\_lX\_default\_active\_minus1; i++) {  
 if (  ViewIdx of the current view component is equal to ViewIdx of RefPicListX[ i ]   
 && (X = = collocated\_from\_l0\_flag || i != collocated\_ref\_idx )   
 && ( NumDdvCandPics ! = 2)) {  
 if (RefPicListX[ i ] is a random access view component) {  
 DdvCandPicsList[ 1 ] = RefPicListX[ i ]  
 NumDdvCandPics =2  
 }  
 else if ( lowestTemporalIDRefs > TemporalID of RefPicListX[ i ])  
 lowestTemporalIDRefs = TemporalID of RefPicListX[ i ]  
 }  
 }  
}~~

1. ~~When NumDdvCandPics is equal to 1, the following applies.~~

~~pocDistance =255  
for ( dir = 0; dir < 2 ; dir++) {  
 X = dir ? collocated\_from\_l0\_flag : (1- collocated\_from\_l0\_flag)  
 for ( i =0 ; i <= num\_ref\_idx\_lX\_default\_active\_minus1; i++) {  
 if ( ViewIdx of the current view component is equal to ViewIdx of RefPicListX[ i ]  
 && (X = = collocated\_from\_l0\_flag || i != collocated\_ref\_idx )  
 && TemporalID of RefPicListX[ i ] = = lowestTemporalIDRefs   
 && (Abs( PicOrderCntVal –~~~~PicOrderCnt( RefPicListX[ i ] ) ) < pocDistance) ){  
 pocDistance = Abs( PicOrderCntVal – PicOrderCnt( RefPicListX[ i ] )  
 Z = X  
 idx = i  
 }  
 }  
}  
if (pocDistance < 255) {  
 DdvCandPicsList[ 1 ] = RefPicListZ[ idx ]  
 NumDdvCandPics =2  
}~~

# Conclusion

In this proposal, an explicit signaling method of the second collocated picture is proposed. Compared with the implicit method in the current HTM, the explicit method provides more encoder flexibility without sacrificing coding efficiency. Moreover, it can be unified with the explicit signaling of the first collocated picture.

# Patent rights declaration (s)

**MediaTek Inc. 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).**

# References

1. G. Tech, K. Wegner Y. Chen, S. Yea, “3D-HEVC Test Model 2,” Document of Joint Collaborative Team on 3D Video Coding Extension Development, JCT3V-B1005, October, 2012.
2. L. Zhang, Y. Chen, M. Karczewicz, “3D-CE5.h related: Disparity vector derivation for multiview video and 3DV,” Document of ISO/IEC JTC 1/SC 29/WG 11, m24937, April, 2012.
3. J. Sung, M. Koo, S. Yea, “3D-CE5.h: Simplification of disparity vector derivation for HEVC-based 3D video coding,” Document of Joint Collaborative Team on 3D Video Coding Extension Development, JCT3V-A0126, July, 2012.
4. J. Kang, Y. Chen, L. Zhang and M. Karczewicz, “3D-CE5.h related: Improvements for disparity vector derivation,” Document of Joint Collaborative Team on 3D Video Coding Extension Development, JCT3V-B0047, October, 2012.
5. HTM-5.0.1, <https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-5.0.1/>
6. D. Rusanovskyy, K. Müller, A. Vetro, “Common test conditions of 3DV Core Experiments,” Document of Joint Collaborative Team on 3D Video Coding Extension Development, JCT3V-A1100, July, 2012.