|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11**  7th Meeting: Geneva, CH, 21-30 November, 2011 | Document: JCTVC-G158 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Undiscardable Leading Pictures for CRA** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Hendry Seungwook Park Byeongmoon Jeon  #221 Yangjae-dong, Seocho-gu,  Seoul 137-130,  Korea | Tel: Email: | +82-10-7738-9288 hendry.hendry@lge.com |
| *Source:* | LG Electronics | | |

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

# Abstract

In current WD 4 of HEVC, decoder shall flush all reference pictures in Decoded Picture Buffer (DPB) prior to decoding the first key picture that follows a CRA picture in decoding order. This contribution shows some use cases when some reference leading pictures should not be flushed out, which are called Undiscardable Leading Pictures (ULPs), prior to decoding the key picture and are allowed to be used as reference for inter prediction for pictures that follow the key picture in order to improve coding efficiency. The contribution proposes some syntax and semantics of new elements to signal ULPs in header of CRA slice. It is reported that by using special input sequences that contains scene change before CRA picture, modified HM-4.0 that implements ULP concept gives gains 0.2 % Y, 0.2% U, 0.2 V for RAHE and 0.3% Y, 0.1% U, 0.2% V for RALC.

# Introduction

There are two types of coded pictures that can be used to allow random access in HEVC encoded bitstream, they are: Instantaneous Decoder Refresh (IDR) picture and Clean Random Access (CRA) picture. While both pictures are coded as intra picture, they command different mechanism of how to treat reference pictures that exist in Decoded Picture Buffer (DPB) and this will affect the referencing mechanism.

Figure 1a and 1b illustrate the different between IDR and CRA pictures. Upon receiving an IDR picture, decoder shall have to flush all reference pictures in DPB by marking them “unused for reference”. Consequently, this shall disallow leading pictures (i.e., pictures that follow IDR / CRA picture in decoding order but precede IDR / CRA in display order), that is picture 25 ~ 31 (picture with POC 25 ~ 31) in Figure 1a, to use pictures in previous GOP as reference for inter prediction because they are simply not available in DPB. On the other hand, when CRA picture is used instead of IDR picture, upon receiving a CRA picture, decoder shall not flush the DPB until next key picture, that is picture 40 (picture with POC 4) in Figure 1b, is received. This will allow the leading pictures to still use pictures in previous GOP as reference pictures. However, because the leading pictures and all other reference pictures in DPB shall be flushed when picture 40 is received, normal pictures (i.e., picture 33 ~ …) shall not be allowed to use the leading pictures as reference pictures.

1. Referencing direction of leading and normal pictures when IDR picture is used
2. Referencing direction of leading and normal pictures when CRA picture is used

Figure 1 – Different between using IDR and CRA pictures

# Undiscardable Leading Pictures

The use of CRA instead of IDR for random access gives coding efficiency improvement because it allows leading pictures to refer to reference pictures in previous GOP with little coding efficiency loss at normal pictures because of the constraint of not using leading pictures as reference pictures.

If normal pictures can use one of more leading pictures as reference pictures, coding efficiency of normal pictures when CRA is used for random access, can be improved. However, for this to be possible, the following requirements must be met:

1. The leading pictures that can be used as reference for normal pictures shall not be marked as “unused for reference” when the first key picture after CRA picture is received.
2. The leading pictures that can be used as reference for normal pictures shall not use other leading picture or other reference picture from previous GOP as their reference pictures.

In this contribution, we refer the leading pictures that are not flushed when the first key picture after CRA is received and can be use as reference picture for normal pictures after the CRA as Undiscardable Leading Pictures (ULPs). Figure 2 illustrates an example of ULP concept in which picture 30 is marked as an ULP. Normally, picture 30 can have 2 forward reference pictures, which are picture 28 and picture 26, and 1 backward reference picture, which is picture 32. However, since an ULP can only refer to other ULPs or a CRA picture, picture 30 shall be encoded without referencing to picture 28 and picture 26. However, by doing this, picture 30 can be safely use as reference picture for normal pictures that follow CRA pictures without breaking random access refresh property.

Figure 2 – Example when picture 30 is marked as ULP

# Use Cases

Under normal condition, severing references for inter prediction can have a serious coding efficiency loss on ULP. However, there are some situations where forward prediction is not required. In this section, we give two scenarios where the use of ULP shall give coding efficiency gain.

## Scene Change within Leading Pictures

If scene change occurs within one GOP before a CRA picture is encoded, reference pictures that precede the picture where the scene change occurs are rarely be selected as reference pictures during inter prediction for pictures that follow the scene change. Thus, when scene change occurs within leading pictures, leading pictures that follow the picture where the scene change first occurs can be chosen as ULP without having coding efficiency loss.

Figure 3 illustrate the scenario when scene change occurs within one GOP before CRA picture. In Figure 3 (a), it is assumed that scene change occurs at the 3rd picture before CRA so that the next picture after that which is used as reference picture are suggested to be set as a ULP. In Figure 3 (a), it is assumed that scene change may occur at either the 4th or 5th picture before CRA so that the 4th picture and the 2nd picture before CRA are suggested to be set as ULPs.

(a) Scene change occurs in leading picture with POC = CRA – 3

(b) Scene change occurs in leading picture with POC = CRA – 4 or CRA - 5

Figure 3 – Example of using ULP when scene change happens within leading pictures

## Encoder Response to Picture Loss within Leading Pictures

In a system with back channel, when picture loss happen to any reference pictures within leading pictures, encoder can decide not to use forward prediction for coding the next pictures to avoid visual quality degradation. The pictures that are coded without forward prediction can then be marked as ULP and be available as reference pictures for normal pictures that follow CRA picture.

For example, suppose that picture with POC CRA – 4 in the example shown in Figure 3(a) is reported to be lost. Then, picture with POC CRA – 2 can be coded without reference to the lost picture and be marked as ULP.

# Syntax and Semantic for Signaling ULP

Table 1 tabulates the syntax of new element to signal ULP pictures in header of CRA slice.

Table 1 – Proposed syntax of slice header for signaling ULP

|  |  |
| --- | --- |
| Slice\_header( ) { | **Desc** |
| **…** |  |
| If (nal\_unit\_type == 4) { |  |
| **number\_of\_undiscardable\_frames** | ue(v) |
| for( i = 0; i < number\_of\_undiscardable\_frames; i++ ) { |  |
| **relative\_poc[ i ]** | ue(v) |
| } |  |
| } |  |
| **…** |  |
| } |  |

**number\_of\_undiscardable\_ frames** indicates the number of undiscardable leading frames that follow the slice in decoding order.

**relative\_poc[i]** indicates the relative POC of the *i*-th undiscardable leading frames. POC of the *i*-th undiscardable leading frames can be computed as follows:

PicOrderCnt(*i*-th ULP) = PicOrderCnt(CurrPic) - relative\_poc[i]

In addition to the syntax and semantic above, the following changes to definition of CRA picture agreed in the 6th JCTVC meeting in Torino [2] is also necessary. The proposed change is highlighted with yellow color.

**clean random access (CRA) picture**: A *coded picture* containing only *I slices* and for which each *slice* of the CRA picture has nal\_unit\_type equal to 4; all *coded pictures* that follow the CRA picture both in *decoding order* and *output order* shall not use *inter prediction* from any *picture* that precedes the CRA picture either in *decoding order* or *output order*; and any *picture* that precedes the CRA picture in *decoding order* also precedes the CRA picture in *output order* except pictures that are marked as undiscardable\_frames in the slice header of the CRA picture.

A flag can be added in SPS to indicate whether ULP frames exist in the bitstream to be decoded. With the addition of such flag, ULP will be checked at slice header of CRA pictures only if the flag indicate that ULPs exist. The following changes are required to accommodate the flag.

Table 2 – Proposed additional flag in SPS

|  |  |
| --- | --- |
| seq\_parameter\_set\_rbsp( ) { | **Desc** |
| **…** |  |
| **no\_ulp\_flag** | f(1) |
| **…** |  |
| } |  |

**no\_ulp\_flag** indicates that there is no ULP exist in the bitstream.

Table 3 – Proposed syntax of slice header for signaling ULP

|  |  |
| --- | --- |
| Slice\_header( ) { | **Desc** |
| **…** |  |
| If (nal\_unit\_type == 4 && !no\_ulp\_flag) { |  |
| **number\_of\_undiscardable\_frames** | ue(v) |
| for( i = 0; i < number\_of\_undiscardable\_frames; i++ ) { |  |
| **relative\_poc[ i ]** | ue(v) |
| } |  |
| } |  |
| **…** |  |
| } |  |

# Simulation Results

The proposed scheme has been implemented on top of HM-4.0 and the performance was measured by simulating the scenario where scene change occurs within leading pictures. For simplicity of the simulation, we modify test sequences that are used in common test condition by interleaving two different test sequences at regular interval to simulate the scene change. For example, test sequence BasketballDrill\_PartyScene\_CRA\_32 is generated by combining class C BasketballDrill and PartyScene at every 32 frames. The resulted sequence has frame order as follow:

* Frame 0 ~ 29 are taken from BasketballDrill sequence
* Frame 30 ~ 61 are taken from PartyScene sequence
* Frame 62 ~ 93 are taken from BasketballDrill sequence
* So forth

Table 4 – Performance of the proposed scheme for RAHE & RALC with anchor HM-4.0 under common test condition except for the test sequences

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access HE** | | | **Random Access LC** | | |
|  | Y | U | V | Y | U | V |
| Class A | -0.1% | -0.2% | -0.2% | -0.4% | 0.0% | -0.2% |
| Class B | -0.2% | -0.2% | -0.3% | -0.3% | -0.2% | -0.2% |
| Class C | -0.2% | -0.1% | 0.0% | -0.2% | -0.1% | -0.1% |
| Class D | -0.2% | -0.2% | -0.3% | -0.2% | -0.1% | -0.1% |
| Class E |  |  |  |  |  |  |
| **Overall** | -0.2% | -0.2% | -0.2% | -0.3% | -0.1% | -0.2% |
|  | -0.2% | -0.2% | -0.2% | -0.3% | -0.1% | -0.2% |
| Enc Time[%] | 100% | | | 101% | | |
| Dec Time[%] | 100% | | | 101% | | |



# Conclusion

This contribution proposes modification to the definition of CRA. In the new definition, when the first key picture after CRA is received, decoder shall flush DPB except reference pictures that are marked as undiscardable leading pictures. It is shown through experiments that ULP can improve coding performance in some situation such as when scene change happens near CRA picture.

We would like to recommend that the JCTVC considers and adopts the proposed ULP concept.

# Patent rights declaration(s)

**LG Electronics 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).**

# Reference

1. JCTVC-F803\_d5, “WD4: Working Draft 4 of High-Efficiency Video Coding,” 6th JCT-VC Meeting, 6th Meeting: Torino, IT, 14-22 July, 2011.
2. JCTVC-F759, “CRA BoG report v2,” 6th JCT-VC Meeting, 6th Meeting: Torino, IT, 14-22 July, 2011.