|  |  |
| --- | --- |
| **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-G166 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **AHG21: Explicit Reference Pictures Signaling with Output Latency Count Scheme** | | |
| *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 the 6th JCTVC meeting, JCTVC-F493 proposed to signal reference pictures that are required by current and future pictures explicitly. This document proposes a scheme that is claimed to improve the scheme proposed in F493. The proposed scheme suggests signaling only reference pictures that are required by current pictures. To assist reference picture marking process, a new syntax element so called Output Latency Count (OLC) is introduced which can tell decoder when each reference picture should be marked as “unused for reference”. It is claimed that the proposed scheme can signal the list of reference pictures more efficiently. Furthermore, it is also claimed that the proposed scheme separates tools for reference picture marking process and error resilience feature so that encoder may have more flexibility to decide whether it wants to use them both. It is reported that while proposed scheme can be use by its own, some part of its concept can be harmonized with signaling scheme developed in AHG21 to address the long-term reference picture issue.

# Introduction

Explicit Reference Picture Signaling (ERPS) was proposed in JCTVC-F493 [2] and currently being further discussed and developed in AHG21. The current ERPS (i.e., the ERPS scheme that was proposed in [2] and currently being further studied in AHG21) can be summarized as follows:

* In each slice header, except header of IDR slices, a set of reference pictures, called reference picture set (RPS), is signal to indicate the list of reference pictures that are required by current picture and future pictures.
* Each reference picture in RPS is signal by its delta POC relative to the POC of the current picture. Furthermore, there is a flag to indicate whether the reference picture is used by current picture.
* After parsing slice header of the current picture, but before decoding process begins, decoder performs reference picture marking process. Reference pictures in Decoded Picture Buffer (DPB) that are not listed in RPS shall be marked as “unused for reference”
* For bitstreams that have structure repetition, RPS can be signaled in PPS.

# Problems with Current ERPS Scheme

## Inefficiency for signaling long term reference pictures

In addition to signaling reference pictures required by current picture, current ERPS scheme also has to signal reference pictures that are required by future pictures. If a reference picture has to be kept for long time in DPB, it is inevitable that current ERPS has to keep on signaling it in every slice header until the reference pictures can finally marked as “unused for reference”.

Furthermore, the longer a reference picture stays in DPB, the bigger its delta POC relative to POC of current picture and thus will require more bits to signal it. Such situation can be clearly observed when long term reference pictures are used.

## Tight coupling of reference picture marking process and error resiliency

The design of the current ERPS has consequence such that reference picture marking process and error resilience feature are tightly coupled because both use RPS in order to operate. For just picture marking process, encoder has no other choice but signaling all required reference pictures. There are cases where encoder might not need error resilience feature provided by current ERPS scheme.

For example, in low delay setting where 3 higher quality and 1 recent reference pictures are used for inter prediction, if buffer management scheme can manage to release reference picture in DPB as soon as it is not needed anymore, the number of reference pictures in DPB at each time shall always be equal to **max\_num\_ref\_frames** + 1 (i.e., current pictures and reference pictures required by current pictures), except for max\_num\_ref\_frame first frames in the bitstreams as shown in Table 1

Table 1 – Ideal DPB management in common test condition for LD

|  |  |  |  |
| --- | --- | --- | --- |
| POC of received pic | DPB content | | |
| Before decoding | After decoding | Released ref pic |
| 0 | - | 0 | - |
| 1 | 0 | 0,1 | - |
| 2 | 0,1 | 0,1,2 | - |
| 3 | 0,1,2 | 0,1,2,3 | - |
| 4 | 0,1,2,3 | 0,1,2,3,4 | - |
| 5 | 0,1,2,3,4 | 0,2,3,4,5 | 1 |
| 6 | 0,2,3,4,5 | 0,3,4,5,6 | 2 |
| 7 | 0,3,4,5,6 | 0,4,5,6,7 | 3 |
| 8 | 0,4,5,6,7 | 0,4,6,7,8 | 5 |
| 9 | 0,4,6,7,8 | 0,4,7,8,9 | 6 |
| 10 | 0,4,7,8,9 | 0,4,8,9,10 | 7 |
| 11 | 0,4,8,9,10 | 0,4,8,10,11 | 9 |
| 12 | 0,4,8,10,11 | 0,4,8,11,12 | 10 |
| 13 | 0,4,8,11,12 | 0,4,8,12,13 | 11 |
| 14 | 0,4,8,12,13 | 4,8,12,13,14 | 0 |
| 15 | 4,8,12,13,14 | 4,8,12,14,15 | 13 |
| 16 | 4,8,12,14,15 | 4,8,12,15,16 | 14 |
| 17 | 4,8,12,15,16 | 4,8,12,16,17 | 15 |
| 18 | 4,8,12,16,17 | 8,12,16,17,18 | 4 |
| … | … | … | … |

Thus, RPS is not needed since picture loss can easily be detected. For example, suppose that picture 9 (i.e., picture with POC 9) is lost and never be received, decoded, and stored in DPB. When picture 10 is received, decoder expects that DPB contains picture 9 (i.e., recent reference picture for picture 10) and picture 0, 4, and 8 (which are high quality reference picture). Note that decoder can easily calculate that picture 10 needs picture 0, 4, and 8 since it knows the GOP size is 8 (picture 0 and 8 are key pictures before picture 10 and picture 4 is the picture in the middle of those key pictures).

In the above situation, RPS is not necessary to be signaled because any picture loss can easily be detected. Therefore, ERPS scheme that mandate encoder to always signal RPS is not necessary and we can save bits by not signaling RPS. All that are required in the above situation is information to tell decoder when a reference picture must be released from DPB.

# Proposed ERPS Scheme

## Design objectives

There are two main objectives in the design of the proposed ERPS scheme:

1. We would like to reduce number of bits required to signal RPS.
2. We would like to separate information that are needed to do reference picture marking process and information that are needed to improve error resiliency.

In order to achieve those objectives, the proposed ERPS scheme is design as follow:

1. The RPS signaled by the proposed ERPS signal contains only reference pictures that are required to decode current picture. This new RPS is used as information to improve error resiliency since any lost can be detected as soon as possible.
2. In first slice header of each picture, a new element called “Output Latency Count” (OLC) is added to inform decoder of when the picture should be release from DPB. This new element is used as information to assist reference picture marking process. OLC is the delta between POC / frame number of current picture and the POC / frame number where current picture should be released from DPB.

By sending only reference pictures that are required by current picture, the proposed ERPS scheme certainly reduces the amount of bits required to signal RPS compared to current ERPS scheme. However, the addition of OLC and necessity to also signal frame number adds overhead bits for the proposed ERPS scheme since current ERPS does not have to signal them.

## Proposed syntax and semantics

Table 2 tabulates the proposed syntax to signal RPS.

Table 2 – Syntax for RPS

|  |  |
| --- | --- |
| ref\_pic\_set( idx ) { | Descriptor |
| **num\_ref\_pics** | ue(v) |
| for( i = 0; i < num\_ref\_pics; i++ ) { |  |
| **sign\_bit\_of\_ref\_pic**[ i ] | ue(v) |
| **delta\_poc\_minus1**[ i ] | u(1) |
| } |  |
| } |  |

**num\_ref\_pics** specifies the number of reference frames signaled in this RPS

**sign\_bit\_of\_ ref\_pic[i]** specifies sign bit of the i-th reference frame. 1 indicates positive sign while 0 indicates otherwise.

**delta\_poc\_minus1[i]** plus 1 specifies the absolute delta POC of the i-th reference frames and the POC of current picture.

For bitstreams with repetitive structures in which reference picture pattern exists for pictures in every GOP, RPS can be signaled in SPS and be referred by their index. Table 3 tabulates the proposed syntax to signal RPS in SPS. In addition to RPS, SPS should also carry some information / flags to indicate the use of OLC and ERPS.

Table 3 – Syntax for signaling RPS in SPS

|  |  |
| --- | --- |
| seq\_parameter\_set\_rbsp( ) { | Descriptor |
| **…** |  |
| **olc\_flag** | f(1) |
| if (olc\_flag) { |  |
| **use\_frame\_num\_flag** | f(1) |
| if (use\_frame\_num\_flag) |  |
| **log2\_max\_frame\_num\_minus4** | ue(v) |
| } |  |
| **erps\_flag** | f(1) |
| if (erps\_flag) { |  |
| **num\_ref\_pic\_sets** | ue(v) |
| for(idx = 0; idx < num\_ref\_pic\_sets; idx++) |  |
| ref\_pic\_set( idx ) | ue(v) |
| **}** |  |
| } |  |
| **…** |  |
| } |  |



**olc \_flag** equals 1 indicates that OLC presents in the header of the first slice of every picture. olc\_flag equals 0 indicates otherwise.

**use\_frame\_num\_flag** indicates that frame\_num is signaled in the slice header.

**log2\_max\_frame\_num\_minus4** specifies the value of the variable MaxFrameNum that is used in frame\_num related derivations as follows:

MaxFrameNum = 2( log2\_max\_frame\_num\_minus4 + 4 )

The value of log2\_max\_frame\_num\_minus4 shall be in the range of 0 to 12, inclusive.

**erps \_flag** specifies whether the reference pictures are signaled explicitly. erps\_flag equals 1 indicates that RPS is signaled explicitly and erps\_flag equals 0 otherwise.

**num\_ref\_pic\_sets** specifies the number of RPS sets signaled in this PPS

Table 4 tabulates the proposed syntax to signal OLC and RPS in slice header

Table 4 – Syntax for signaling RPS in Slice Header

|  |  |
| --- | --- |
| slice\_header ( ) { | Descriptor |
| **…** |  |
| if (nal\_ref\_flag == 1 && olc\_flag) { |  |
| if (use\_frame\_num\_flag) { |  |
| **frame\_num** | u(v) |
| } |  |
| **olc\_minus2** | ue(v) |
| } |  |
| If ((slice\_type == P || slice\_type == B) && erps\_flag == 1) { |  |
| **ref\_pic\_set\_idx\_plus1** | ue(v) |
| if (ref\_pic\_set\_idx\_plus1 > 0) { //not using RPS in SPS |  |
| ref\_pic\_set( idx - 1) | ue(v) |
| **}** |  |
| } |  |
| **…** |  |
| } |  |



**frame\_num** as semantic specified in WD 4 of HEVC.

**olc\_minus2** plus 2specifies the relative indicates how long the current picture shall stay in the DPB before it can be marked as “unused for reference”. For bitstream with low delay settings, OLC is relative to POC, otherwise, it is relative to frame\_num.

If (nal\_ref\_flag == 0)

OutputLatencyCount = 0

Mark current picture as “unused for reference”

Else if (nal\_ref\_flag == 1 && !use\_frame\_num\_flag)

OutputLatencyCount = olc\_minus2 + 2 + PicOrderCnt(CurrPic)

Mark current picture as “used for reference”

Else

OutputLatencyCount = olc\_minus2 + 2 + frame\_num(CurrPic)

Mark current picture as “used for reference”

Reference Frame Marking Process

This initialization process is invoked after decoding the first slice header of every frame and updating OutputLatencyCount of ExistingReferenceFrame in DPB.

If (!use\_frame\_num\_flag) {

For i = 0 to number of reference frames in DPB, do:

If (reference\_frame[i] != CurrPic frame && OutputLatencyCount[i] > 0   
&& (OutputLatencyCount[i] – frame\_num(CurrPic)) <= 0)

Mark reference\_frame[i] as “unused for reference”

}

Else

{

For i = 0 to number of reference frames in DPB, do:

If (reference\_frame[i] != CurrPic frame && OutputLatencyCount[i] > 0   
&& (OutputLatencyCount[i] – PicOrderCnt(CurrPic)) <= 0)

Mark reference\_frame[i] as “unused for reference”

}

**ref\_pic\_set\_idx\_plus1** minus 1specifies the index of RPS stored in SPS. Ref\_pic\_set\_idx\_plus1 equals 0 indicates that RPS is signaled in current slice header.

## Addressing Long-term Reference Picture Signaling Issue

The syntaxes and semantics proposed in Section 3.2 can be extended if long-term reference pictures (LTRPs) concept is used.

To accommodate LTRPs, the following syntaxes and semantics are proposed.

1. A flag called “use\_long\_term\_reference\_flag” should be signaled in SPS. This is similar to Microsoft’s proposal – JCTVC-G788.

Table 5 – Syntax for signaling RPS in SPS

|  |  |
| --- | --- |
| seq\_parameter\_set\_rbsp( ) { | Descriptor |
| **…** |  |
| **use\_long\_term\_reference\_flag** | f(1) |
| **…** |  |
| } |  |

**use\_long\_term\_reference\_flag** equals 0 indicates that LTRP does not present in the bitstream. use\_long\_term\_reference\_flag equal 1 indicates that LTRP may present in the bitstream.

1. A flag called “LTRP\_flag” should be signaled in Slice Header to indicate whether current picture shall be used as LTRP.
2. For each LTRP, a unique id (i.e., LTRP\_Id) shall be given. Reference to LTRP in slice header shall use LTRP\_Id, instead of delta POC.
3. OLC for LTRP is signaled in special way to reduce the number of required bits.

Table 6 – Syntax for signaling RPS in SPS

|  |  |
| --- | --- |
| slice\_header ( ) { | Descriptor |
| **…** |  |
| if (nal\_ref\_flag == 1 && use\_long\_term\_reference\_flag) { |  |
| **LTRP\_flag** | f(1) |
| If (LTRP\_flag == 1) { |  |
| **LTRP\_Id** | ue(v) |
| **log2\_olc\_msb\_minus4** | ue(v) |
| **olc\_cnt** | u(v) |
| } |  |
| } |  |
| **…** |  |
| } |  |

**LTRP\_flag** equals 0 indicates that current picture is not a LTRP. LTRP\_flag equal 1 indicates otherwise.

**LTRP\_Id** indicate the LTRP identification number of the current picture. LTRP\_Id shall be unique in decoded picture buffer (DPB). If other LTRP with same LTRP\_Id has already existed in DPB, prior to storing current picture into DPB, the pre-existing LTPR with same LTRP\_Id shall be first marked as “unused for reference”

**log2\_olc\_msb\_minus4** plus 4 indicate the number of bits assigned for OLC\_msb

**olc\_cnt** indicates the output latency count of the LTRP. The number of bit assigned to signal olc\_cnt is log2\_olc\_msb\_minus4 + 4.

OLC of current picture is calculated as follows:

OLC\_msb = 2log2\_OLC\_msb\_minus4 + 4

OLC (CurrPic) = OLC\_msb + OLC\_cnt

A LTRP shall be marked as “unused for reference” if the following condition is met:

POC (LTRP) + OLC (LTRP) <= POC (CurrPic)

The signaling scheme for LTRP can easily be integrated into OLC and ERPS scheme (i.e., for short-term reference pictures -- STRP) proposed in Section 3.2. However, if it is desired, it can also be harmonized with ERPS scheme being developed in AHG21.

### Integration of LTRP with ERPS scheme proposed in 3.2

The signaling scheme for LTRP can be integrated into ERPS scheme proposed in 3.2 by modifying the syntax of slice header shown in Table 4 as follows:

Table 7 – ERPS for signaling long-term and short-term reference pictures

|  |  |
| --- | --- |
| slice\_header ( ) { | Descriptor |
| **…** |  |
| if (nal\_ref\_flag == 1 && use\_long\_term\_reference\_flag) { |  |
| **LTRP\_flag** | f(1) |
| If (LTRP\_flag == 1) { |  |
| **LTRP\_Id** | ue(v) |
| **log2\_olc\_msb\_minus4** | ue(v) |
| **olc\_cnt** | u(v) |
| } else if (nal\_ref\_flag == 1 && olc\_flag) { |  |
| if (use\_frame\_num\_flag) { |  |
| **frame\_num** | u(v) |
| } |  |
| **olc\_minus2** | ue(v) |
| } |  |
| } else if (nal\_ref\_flag == 1 && olc\_flag) { |  |
| if (use\_frame\_num\_flag) { |  |
| **frame\_num** | u(v) |
| } |  |
| **olc\_minus2** | ue(v) |
| **}** |  |
| If ((slice\_type == P || slice\_type == B) && erps\_flag == 1) { |  |
| **ref\_pic\_set\_idx\_plus1** | ue(v) |
| if (ref\_pic\_set\_idx\_plus1 > 0) { //not using RPS in SPS |  |
| ref\_pic\_set( idx - 1) |  |
| } |  |
| } |  |
| If (use\_long\_term\_reference\_flag == 1) { |  |
| **num\_used\_long\_term\_pics** | ue(v) |
| If (num\_used\_long\_term\_pics > 0) { |  |
| for (i = 0; i < num\_used\_long\_term\_pics; i++) { |  |
| LTRP\_Id | ue(v) |
| } |  |
| } |  |
| } |  |
| **…** |  |
| } |  |

**num\_used\_long\_term\_pics** indicates the number of LTRPs required by current picture for its inter prediction

### Integration of LTRP with ERPS scheme in AHG21

To harmonize the proposed LTRP signaling scheme with current ERPS scheme in AHG21, the following modifications are proposed:

1. Reference picture marking process for LTRPs should be different from for STRPs. Reference pictures marking process for STRPs uses current mechanism described in the WD of AHG21 whereas marking process for LTRP use the mechanism described in Section 3.3, that is an LTRP shall be marked as “unused for reference” if:
   1. There is a new LTRP with the same LTRP\_Id, or
   2. POC (LTRP) + OLC (LTRP) <= POC (CurrPic)
2. Instead of signaling **num\_long\_term\_pics**, we propose to signal **num\_used\_long\_term\_pics**.
3. Reference to LTRP should not use delta POC because as delta POC gets bigger as LTRP gets farther from current picture, which can be costly to signal it. Instead of delta POC, we propose to refer LTRP by its Id.

With the above modification, the slice header should be modified as follow:

Table 8 – harmonized version of proposed LTRP and STRP in AHG21

|  |  |
| --- | --- |
| slice\_header( ) { | Descriptor |
| **…** |  |
| if (nal\_ref\_flag == 1 && use\_long\_term\_reference\_flag) { |  |
| **LTRP\_flag** | f(1) |
| If (LTRP\_flag == 1) { |  |
| **LTRP\_Id** | ue(v) |
| **log2\_olc\_msb\_minus4** | ue(v) |
| **olc\_cnt** | ue(v) |
| } |  |
| } |  |
| if( IdrPicFlag ) { |  |
| **idr\_pic\_id** | ue(v) |
| **no\_output\_of\_prior\_pics\_flag** | u(1) |
| } |  |
| else { |  |
| **pic\_order\_cnt\_lsb** | u(v) |
| **short\_term\_ref\_pic\_set\_pps\_flag** | u(1) |
| if( !short\_term\_ref\_pic\_set\_pps\_flag ) |  |
| short\_term\_ref\_pic\_set( num\_short\_term\_ref\_pic\_sets ) |  |
| else |  |
| **short\_term\_ref\_pic\_set\_idx** | u(v) |
| if( long\_term\_ref\_pics\_present\_flag ) { |  |
| **num\_used\_long\_term\_pics** | ue(v) |
| for( i = 0; i < num\_used\_long\_term\_pics; i++ ) { |  |
| LTRP\_Id | ue(v) |
| } |  |
| } |  |
| } |  |
| **…** |  |
| } |  |

# Discussion

## Proposed ERPS Scheme vs. Picture Loss / Temporal Layer Removal

Since the proposed ERPS scheme separate the information to assist reference picture marking process and information to improve error resilience, the robustness of the proposed ERPS scheme to handle reference picture marking process might in questioned.

The following examples show the robustness of the OLC in the proposed ERPS scheme against picture loss and temporal layer removal.

Example 1 – Normal decoding, No picture loss & no temporal layer removal

Example 1 gives an example of decoding and reference frame marking process when all frames are received normally (no extracted temporal layer, no frame loss). Figure 1 illustrates the structure of bitstream for the example. In brief, the marking process is as follows:

* Frame 20 (i.e., frame with POC 20) can be released from buffer before decoding of frame with frame number 27 (that is 20 + 5 + 2) starts.
* Frame 21 can be released from buffer before decoding of frame 24 starts.
* Frame 22 can be released from buffer before decoding of frame 24 starts.
* Frame 23 can be released from buffer before decoding of frame 26 starts.
* Frame 24 can be released from buffer before decoding of frame 31 starts.
* Frame 25 can be released from buffer before decoding of frame 28 starts.
* Frame 26 can be released from buffer before decoding of frame 28 starts.
* And so forth…

Figure 1 – Bitstream structure for example 1, 2, and 3

Example 2 – Picture loss occurs

By using the structure shown in Figure 1, assume that:

1. Frame 23 is lost: no action is needed. Reference picture marking process of other frames will be be affected. The lost of frame 23 will be detected when frame 25 is received.
2. Frame 24 is lost: Frame 21 and 22 that supposed to be released from DPB when frame 24 is received can be released from DPB as soon as the next frame (i.e., frame 25) is received. The lost of frame 24 will be detected when frame 25 is received.

Example 3 – Temporal layer removal occurs

By using the structure shown in Figure 1, assume that the highest temporal layer is removed so that frame 21, 23, 25, and 27 are not received by decoder. Both decoding process and reference picture marking process will not be affected

## OLC with frame\_num vs. OLC with POC

Reference picture marking process can use OLC to decide status of reference pictures by comparing their OLC to either frame number of POC. Both uses (i.e., OLC with frame number and OLC with POC) have their own trade-off.

The following two sections provide use case examples when using OLC with frame number and OLC with POC in random access setting – GOP size 8, intra period 32.

### OLC with frame number

Table 9 tabulates received frames, DPB state before and after decoding each frame, and number of overhead bits to signal to accommodate the proposed scheme.

Table 9 – Overhead bits from ERPS with proposed scheme   
where OLC is used together with frame number

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Frame Number** | **POC** | **OLC** | **DPB content** | | | **Overhead bits** | | |
| **Before decoding** | **Signaled Ref Pics** | **After decoding** | **Frame Number** | **OLC** | **LC** |
| 0 | 0 | 4 |  |  | 0 | 4 | 5 | 0 |
| 1 | 8 | 7 | 0 | [LC 0] | 0,8 | 4 | 7 | 8 |
| 2 | 4 | 2 | 0,8 | [LC 0 8 ] | 0,8,4 | 4 | 5 | 12 |
| 3 | 2 | 1 | 0,8,4 | [LC 0 4 8 ] | 0,8,4,2 | 4 | 3 | 14 |
| 4 | 6 | 4 | 0,8,4,2 | [LC 4 8 2 ] | 0,8,4,2,6 | 4 | 5 | 12 |
| 5 | 1 | - | 0,8,4,2,6 | [LC 0 2 4 ] | 0,8,4,2,6 | 4 | 0 | 12 |
| 5 | 3 | - | 0,8,4,2,6 | [LC 2 4 0 6 ] | 0,8,4,2,6 | 4 | 0 | 16 |
| 5 | 5 | - | 0,8,4,2,6 | [LC 4 6 2 8 ] | 0,8,4,2,6 | 4 | 0 | 16 |
| 5 | 7 | - | 0,8,4,2,6 | [LC 6 8 4 ] | 0,8,4,2,6 | 4 | 0 | 12 |
| 5 | 16 | 7 | 0,8,4,2,6 | [LC 8 6 4 2 ] | 0,8,4,2,6,16 | 4 | 7 | 20 |
| 6 | 12 | 2 | 0,8,4,2,6,16 | [LC 8 16 6 ] | 8,6,16 | 4 | 3 | 16 |
| 7 | 10 | 1 | 8,6,16 | [LC 8 12 6 16 ] | 8,6,16,12 | 4 | 3 | 18 |
| 8 | 14 | 4 | 8,6,16,12 | [LC 12 16 10 ] | 8,6,16,12,10 | 4 | 5 | 12 |
| 9 | 9 | - | 8,6,16,12,10 | [LC 8 10 6 12 ] | 8,6,16,12,10,14 | 4 | 0 | 16 |
| 9 | 11 | - | 8,6,16,12,10,14 | [LC 10 12 8 14 ] | 8,6,16,12,10,14 | 4 | 0 | 16 |
| 9 | 13 | - | 8,6,16,12,10,14 | [LC 12 14 10 16 ] | 8,6,16,12,10,14 | 4 | 0 | 16 |
| 9 | 15 | - | 8,6,16,12,10,14 | [LC 14 16 12 ] | 8,6,16,12,10,14 | 4 | 0 | 12 |
| 9 | 24 | 7 | 8,6,16,12,10,14 | [LC 16 14 12 10 ] | 8,6,16,12,10,14,24 | 4 | 7 | 20 |
| 10 | 20 | 2 | 8,6,16,12,10,14,24 | [LC 16 24 14 ] | 16,14,24,20 | 4 | 3 | 16 |
| 11 | 18 | 1 | 16,14,24,20 | [LC 16 20 14 24 ] | 16,14,24,20,18 | 4 | 3 | 18 |
| 12 | 22 | 4 | 16,14,24,20,18 | [LC 20 24 18 ] | 16,14,24,20,18,22 | 4 | 5 | 12 |
| 13 | 17 | - | 16,14,24,20,18,22 | [LC 16 18 14 20 ] | 16,14,24,20,18,22 | 4 | 0 | 16 |
| 13 | 19 | - | 16,14,24,20,18,22 | [LC 18 20 16 22 ] | 16,14,24,20,18,22 | 4 | 0 | 16 |
| 13 | 21 | - | 16,14,24,20,18,22 | [LC 20 22 18 24 ] | 16,14,24,20,18,22 | 4 | 0 | 16 |
| 13 | 23 | - | 16,14,24,20,18,22 | [LC 22 24 20 ] | 16,14,24,20,18,22 | 4 | 0 | 12 |
| **…** | | | | | | | | |
|  | | | | | Subtotal bits | 100 | 61 | 354 |
| **Total bits** | **515** |  |  |

### OLC with POC

Table 10 tabulates received frames, DPB state before and after decoding each frame, and number of overhead bits to signal to accommodate the proposed scheme.

Table 10 – Overhead bits from ERPS with proposed scheme where OLC is used together with POC

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **POC** | **OLC** | **DPB content** | | |  | |
| **Before decoding** | **Signaled Ref Pics** | **After decoding** | **OLC** | **LC** |
| 0 | 7 |  |  | 0 | 7 | 0 |
| 8 | 7 | 0 | [LC 0] | 0,8 | 7 | 8 |
| 4 | 11 | 0,8 | [LC 0 8 ] | 0,8,4 | 7 | 12 |
| 2 | 13 | 0,8,4 | [LC 0 4 8 ] | 0,8,4,2 | 7 | 14 |
| 6 | 9 | 0,8,4,2 | [LC 4 8 2 ] | 0,8,4,2,6 | 7 | 12 |
| 1 | - | 0,8,4,2,6 | [LC 0 2 4 ] | 0,8,4,2,6 | 0 | 12 |
| 3 | - | 0,8,4,2,6 | [LC 2 4 0 6 ] | 0,8,4,2,6 | 0 | 16 |
| 5 | - | 0,8,4,2,6 | [LC 4 6 2 8 ] | 0,8,4,2,6 | 0 | 16 |
| 7 | - | 0,8,4,2,6 | [LC 6 8 4 ] | 0,8,4,2,6 | 0 | 12 |
| 16 | 7 | 0,8,4,2,6 | [LC 8 6 4 2 ] | 8,4,2,6,16 | 7 | 20 |
| 12 | 11 | 8,4,2,6,16 | [LC 8 16 6 ] | 8,4,2,6,16 | 7 | 16 |
| 10 | 13 | 8,4,2,6,16 | [LC 8 12 6 16 ] | 8,4,2,6,16,12 | 7 | 18 |
| 14 | 9 | 8,4,2,6,16,12 | [LC 12 16 10 ] | 8,4,2,6,16,12,10 | 7 | 12 |
| 9 | - | 8,4,2,6,16,12,10 | [LC 8 10 6 12 ] | 8,4,2,6,16,12,10,14 | 0 | 16 |
| 11 | - | 8,4,2,6,16,12,10,14 | [LC 10 12 8 14 ] | 8,4,2,6,16,12,10,14 | 0 | 16 |
| 13 | - | 8,4,2,6,16,12,10,14 | [LC 12 14 10 16 ] | 8,4,2,6,16,12,10,14 | 0 | 16 |
| 15 | - | 8,4,2,6,16,12,10,14 | [LC 14 16 12 ] | 8,4,2,6,16,12,10,14 | 0 | 12 |
| 24 | 7 | 8,4,2,6,16,12,10,14 | [LC 16 14 12 10 ] | 16,12,10,14,24 | 7 | 20 |
| 20 | 11 | 6,16,12,10,14,24 | [LC 16 24 14 ] | 16,12,10,14,24,20 | 7 | 16 |
| 18 | 13 | 16,12,10,14,24,20 | [LC 16 20 14 24 ] | 16,12,10,14,24,20,18 | 7 | 18 |
| 22 | 9 | 16,12,10,14,24,20,18 | [LC 20 24 18 ] | 16,12,10,14,24,20,18,22 | 7 | 12 |
| 17 | - | 16,12,10,14,24,20,18,22 | [LC 16 18 14 20 ] | 16,12,10,14,24,20,18,22 | 0 | 16 |
| 19 | - | 16,12,10,14,24,20,18,22 | [LC 18 20 16 22 ] | 16,12,10,14,24,20,18,22 | 0 | 16 |
| 21 | - | 16,12,10,14,24,20,18,22 | [LC 20 22 18 24 ] | 16,12,10,14,24,20,18,22 | 0 | 16 |
| 23 | - | 16,12,10,14,24,20,18,22 | [LC 22 24 20 ] | 16,12,10,14,24,20,18,22 | 0 | 12 |
|  | | | | | | |
|  |  |  |  | Subtotal bits | 91 | 354 |
|  |  |  |  | **Total bits** | **445** |  |

The trade-off between using OLC with frame number and with POC can be summarized as follows:

* Required size of DPB is smaller when OLC is used with frame number. This is because picture that are no required for reference pictures can be marked “unused for reference” earlier.
* Overhead bit for ERPS is smaller when OLC is used with POC. This is because there is no need to signal frame number anymore.

# Simulation Results

The proposed ERPS scheme has been implemented on top of HM-4.0 and the performance was measured under common test condition for RAHE, RALC, LBHE, and LBLC configuration. The results are shown in Table 11 ~ 13. The simulation for ERPS is conducted by signaling RPS in slice header, not in SPS. Furthermore, RPS is only signaled in random access settings (RAHE and RALC) while for low delay settings (LBHE and LBLC), we only signal OLC.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Table 11 – Simulation result for RAHE & RALC where OLC is used together with frame number | **Random Access HE** | | | **Random Access LC** | | |
|  | Y | U | V | Y | U | V |
| Class A | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Class B | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Class C | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% |
| Class D | 0.3% | 0.2% | 0.2% | 0.2% | 0.2% | 0.2% |
| Class E |  |  |  |  |  |  |
| **Overall** | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% |
|  | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% |
| Enc Time[%] | 100% | | | 100% | | |
| Dec Time[%] | 93% | | | 94% | | |

Table 12 – Simulation result for RAHE & RALC where OLC is used together with POC

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Random Access HE** | | | **Random Access LC** | | |
|  | Y | U | V | Y | U | V |
| Class A | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Class B | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Class C | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% |
| Class D | 0.2% | 0.2% | 0.2% | 0.2% | 0.2% | 0.2% |
| Class E |  |  |  |  |  |  |
| **Overall** | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% |
|  | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% | 0.1% |
| Enc Time[%] | 100% | | | 100% | | |
| Dec Time[%] | 96% | | | 99% | | |

for random access cases It is also worth to note that the use of frame number does not add any significant overhead as it is shown that with or without frame number, the average bit-rate increase is still the same.

The signaling scheme of the proposed ERPS without frame number, whose results shown in Table 12, is slightly more efficient than that of the current ERPS in AHG21. This can be easily observed by comparing performance loss due to overhead bit for signaling for sequences in class C and D. Please see attached spreadsheets JCTVC-G166\_OLC\_WithoutFrameNum.xlsx and JCTVC-G166\_HM4.0-Dev.vs.ERPSinAHG21.xlsx

For low delay case where only OLC is signaled, there is no change in the average output bit-rate as demonstrated by the result shown in Table 13.

Table 13 – Simulation result for LBHE & LBLC – Anchor HM-4.0 common test condition

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Low delay B HE** | | | **Low delay B LC** | | |
|  | Y | U | V | Y | U | V |
| Class A |  |  |  |  |  |  |
| Class B | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Class C | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Class D | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Class E | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| **Overall** | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
|  | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Enc Time[%] | 100% | | | 100% | | |
| Dec Time[%] | 94% | | | 100% | | |



# Conclusion

This contribution proposes an explicit reference picture signaling (ERPS) scheme for reference picture management of HEVC. The proposed scheme can be considered as a modified version of the current ERPS that is being discussed and developed in AHG21. While the proposed scheme can be use by its own, some part of its concept can be harmonized with signaling scheme developed in AHG21 to address the long-term reference picture issue.

We would like to recommend that the JCTVC considers the proposed ERPS scheme and further discussed it in the scope of AHG21.

# 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-F493, “Absolute signaling of reference pictures,” 6th JCT-VC Meeting, 6th Meeting: Torino, IT, 14-22 July, 2011.