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-G715

Title: **AHG18/21: Absolute signaling for resolution switching**

Status: Input Document to JCT-VC

Purpose: Proposal

Author(s) or Contact(s): Kiran Misra, Sachin Deshpande, Louis Kerofsky, Andrew Segall
5750 Pacific Rim Blvd,
Camas, WA 98607,
USA.

Tel: +1 360 817 8451
Email: misrak@sharplabs.com
sdeshpande@sharplabs.com
lkerofsky@sharplabs.com
asegall@sharplabs.com

Source: Sharp Labs of America, Inc.

Abstract

This document describes a technique which enables signaling of reference picture resolutions to be maintained in the decoded picture buffer.

1 Introduction

In JCTVC-F158 [1] a method for changing the picture resolution without causing an instantaneous decoder refresh (IDR) or new sequence parameter set (SPS) to be sent was proposed. One scenario where switching to a lower resolution is desirable is when the available network bandwidth reduces significantly and it is desirable to have a lower bitrate bitstream. To study and compare other network adaptation approaches such as coarser quantization to resolution switching an ad-hoc group AHG18 [2] was formed at the 6th JCT-VC meeting [3].

JCTVC-F493 [5] proposed absolute signaling of reference pictures which supports error robust management of reference pictures using Reference Picture Sets. To study possible improvements based on [5] to the reference picture list construction process an ad-hoc group AHG21 [6] was formed at the 6th JCT-VC meeting [3]. One of the outputs of AHG21 has been a candidate working draft specification [7] for signaling reference pictures.

The current document builds on top of the candidate working draft specification in [7] to enable signaling of reference picture resolutions to be maintained in the decoded picture buffer.

2 Proposed change

If a stream switches to a lower resolution without an IDR or signaling a new SPS then the decoded picture buffer may contain pictures having different resolutions. An encoder may then signal to the decoder that the reference pictures in the decoded picture buffer (DPB) are maintained at:

a. The lower resolution only

b. The higher resolution only

c. Both the higher and lower resolution

Some example use-cases:

TCL-1013
TTE Technology Inc. d/b/a TCL North America v. InterDigital, Inc. IPR2026-00417
AMAZON-1013
Amazon v. InterDigital, IPR2026-00195

Option (a) is useful if it is known that a reference picture will only be used for motion prediction at the lower resolution (example illustrated in Figure 1(a)).

Option (b) is useful when the stream switches back to higher resolution and it is desired to use the long term picture stored in the DPB for motion prediction (example illustrated in Figure 1(b)).

Option (c) is useful when a lower resolution version of the reference picture is used in motion prediction for lower resolution pictures. Additionally, the higher resolution version of the reference picture is used for motion prediction once the bitstream switches back to the original resolution (example illustrated in Figure 1(c))

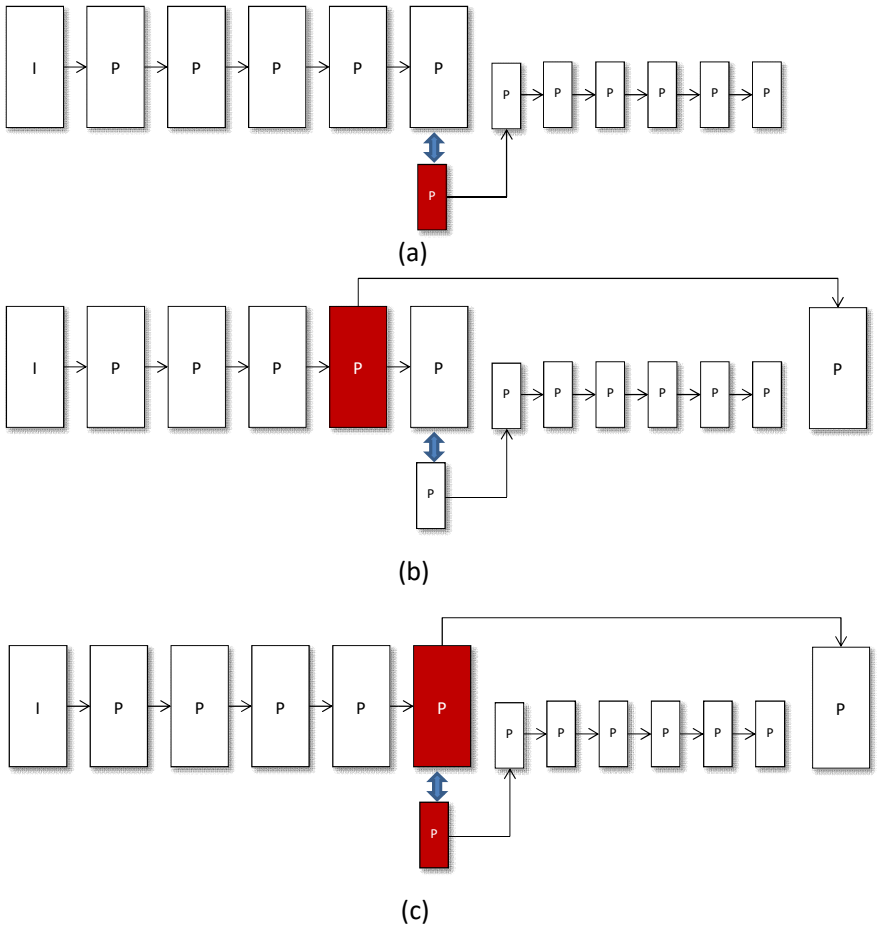

Figure 1: Examples illustrating different picture resolutions to be maintained for motion prediction (a) Only a lower resolution picture should be stored in DPB (b) Only the higher resolution picture should be stored in DPB (c) Both higher and lower resolution picture should be stored in DPB .

3 Syntax and Semantics

Based on the syntax described in [7] the following syntax may be defined to signal which resolution of reference pictures be maintained in the DPB.

ref_pic_set(idx) {	**Descriptor**
num_negative_pics	ue(v)
num_positive_pics	ue(v)
num_longterm_pics	ue(v)
for(i = 0; i < num_negative_pics; i++) {	
delta_poc_s0_minus1[i]	ue(v)
used_by_curr_pic_s0_flag[i]	u(1)
if ((refPicS0Width[i]!=curPicWidth] || refPicS0Height[i]!=curPicHeight) && adaptive_res_coding_flag)	
ref_pic_s0_resolution_idx[i]	ue(v)
}	
for(i = 0; i < num_positive_pics; i++) {	
delta_poc_s1_minus1[i]	ue(v)
used_by_curr_pic_s1_flag[i]	u(1)
if ((refPicS1Width[i]!=curPicWidth || refPicS1Height[i]!=curPicHeight) && adaptive_res_coding_flag)	
ref_pic_s1_resolution_idx[i]	ue(v)
}	
for(i = 0; i < num_longterm_pics; i++) {	
longterm_poc [i]	ue(v)
used_by_curr_pic_lt_flag[i]	u(1)
if ((refPicS1Width[i]!=curPicWidth || refPicS1Height[i]!=curPicHeight) && adaptive_res_coding_flag)	
ref_pic_longterm_resolution_idx[i]	ue(v)
}	
}	

adaptive_res_coding_flag is as defined in [4]. If **adaptive_res_coding_flag** (which is signaled in SPS) equals 0 then the sequence does not undergo resolution switching. If **adaptive_res_coding_flag** equals to 1 then the sequence undergoes resolution switching.

refPicS0Height[i], refPicS0Width[i] represents the original height and width of the i-th reference picture corresponding to negative deltaPOC.

refPicS1Height[i], refPicS1Width[i] represents the original height and width of the i-th reference picture corresponding to positive deltaPOC.

curPicHeight, curPicWidth represents the height and width of the current picture being decoded.

ref_pic_s0_resolution_idx[i] equals to 0 specifies that the reference picture is scaled to make its resolution match the resolution of the current picture being decoded. **ref_pic_s0_resolution_idx[i]** equals to 1 specifies that the reference picture is kept at the reference pictures original resolution. **ref_pic_s0_resolution_idx[i]** equals to 2 specifies that the reference picture resolution is maintained at both at its original resolution and the resolution of the current picture. **ref_pic_s0_resolution_idx[i]** is inferred to be 0 when not present.

ref_pic_s1_resolution_idx[i] equals to 0 specifies that the reference picture is scaled to make its resolution match the resolution of the current picture being decoded. **ref_pic_s1_resolution_idx[i]** equals to 1 specifies that the reference picture is kept at the reference pictures original resolution. **ref_pic_s1_resolution_idx[i]** equals to 2 specifies that the reference picture resolution is maintained at both at its original resolution and the resolution of the current picture. **ref_pic_s1_resolution_idx[i]** is inferred to be 0 when not present.

num_longterm_pics specifies the number of long term pictures specified by syntax elements longterm_poc[i], used_by_curr_pic_pic_lt_flag[i]. The value of num_longterm_pics shall be in the range of 0 to max_num_ref_frames – num_negative_pics,- num_positive_pics inclusive.

longterm_poc [i] specifies the picture order count of the i-th long term picture being referenced. The longterm_poc [i] is the sum of POC MSB and POC LSB for the long term picture being referenced.

used_by_curr_pic_lt_flag[i] equal to 0 specifies that the i-th long term picture is not used for reference by the current picture. **used_by_curr_pic_lt_flag[i]** equal to 1 specifies that the i-th long term picture is used for reference by the current picture.

ref_pic_longterm_resolution_idx[i] equals to 0 specifies that the long term reference picture is scaled to make its resolution match the resolution of the current picture being decoded. **ref_pic_longterm_resolution_idx[i]** equals to 1 specifies that the long term reference picture is kept at the long term reference pictures original resolution. **ref_pic_longterm_resolution_idx[i]** equals to 2 specifies that the long term reference picture resolution is maintained at both at its original resolution and the resolution of the current picture. **ref_pic_longterm_resolution_idx[i]** is inferred to be 0 when not present.

4 Conclusion

This document proposes a technique which enables signaling of reference picture resolutions to be maintained in the decoded picture buffer. The proposed syntax and semantics is based on the candidate specification working draft document [7] output by AHG21. We propose to include the approach in any HEVC design that includes resolution switching and reference picture set concept.

5 Patent rights declaration(s)

Sharp 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).

6 References

[1] T. Davies (Cisco) “Resolution switching for coding efficiency and resilience,” JCTVC-F158/m20577, 6th Meeting: Turin, IT, 14-22 July 2011.

[2] “JCT-VC AHG report: Resolution adaption (AHG 18),” JCTVC-G018/m21762, 7th Meeting: Geneva, CH, 21-30 November, 2011.

[3] 6th JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 meeting: Turin, IT, 14-22 July 2011.

[4] T. Davies (Cisco), P. Topiwala (FastVDO), “AHG18: Adaptive Resolution Coding (ARC) “, JCTVC-G264/m21824, 7th Meeting: Geneva, CH, 21-30 November, 2011.

[5] R. Sjöberg, J. Samuelsson (Ericsson), "Absolute signaling of reference pictures," JCTVC-F493/m20923 6th Meeting: Torino, IT, 14-22 July, 2011.

[6] D. Flynn, R. Sjöberg (AHG chairs), Y. Chen, T.K. Tan, W. Wan, Y.-K. Wang (vice chairs), " JCT-VC AHG report: Reference picture buffering and list construction (AHG21)" JCTVC-G021 7th Meeting: Geneva, CH, 21-30 November, 2011.

[7] ftp://ftp.hhi.de/ahg21/JCTVC-F803_d5_RPS_r1.doc Candidate Working Draft text by AHG 21