

JCTVC-N0217/JCT3V-E0097 ON SHVC HIGH LEVEL SYNTAX

Yong He, Xiaoyu Xiu, Yan Ye, Yuwen He

max_one_active_ref_layer_flag

- Propose max_num_active_ref_layers_minus1 to replace max_one_active_ref_layer_flag
- Benefits
 - Maintain same bit cost when only one active ref layer is allowed (1 bit)
 - Inform the codec/middle box the number of maximum ref layers to be supported
 - Flexible for future scalable profile/levels constrains

Proposed syntax

vps_extension() {	Descriptor
while(!byte_aligned())	
vps_extension_byte_alignment_reserved_one_bit	u(1)
...	
max_one_active_ref_layer_flag	u(1)
max_num_active_ref_layers_minus1	ue(v)
...	
}	

slice_segment_header() {	Descriptor
...	
if(nuh_layer_id > 0 && NumDirectRefLayers[nuh_layer_id] > 0) {	
inter_layer_pred_enabled_flag	u(1)
if(inter_layer_pred_enabled_flag && NumDirectRefLayers[nuh_layer_id] > 1) {	
if(max_one_active_ref_layer_flag max_num_active_ref_layers_minus1)	
num_inter_layer_ref_pics_minus1	u(v)
for(i = 0; i < NumActiveRefLayerPics; i++)	
inter_layer_pred_layer_idc[i]	u(v)
}	
}	
...	
}	

Proposed semantic

- **max_num_active_ref_layers_minus1** plus 1 specifies the maximum number of pictures used for inter-layer prediction for each picture in the CVS. The value of **max_num_active_ref_layers_minus1** shall be in the range of 0 to **NumDirectRefLayers[nuh_layer_id]**, exclusively. **max_num_active_ref_layers_minus1** equal to 0 specifies that only one picture may be used for inter-layer prediction for each picture in the CVS. When **max_num_active_ref_layers_minus1** is not present, it is inferred to be equal to 0. When **max_num_active_ref_layers_minus1** is equal to 0, the syntax element **num_inter_layer_ref_pics_minus1** may not be signaled in the slice header.

- The derivation of **NumActiveRefLayerPics** as specified in SHVC WD2 G.7.4.7.1 can be modified as follows:

The variable **NumActiveRefLayerPics** is derived as follows:

```
if( nuh_layer_id == 0 || NumDirectRefLayers[ nuh_layer_id ] == 0 ||  
!inter_layer_pred_enabled_flag )  
    NumActiveRefLayerPics = 0  
else if( !max_one_active_ref_layer_flag !max_num_active_ref_layers_minus1 ||  
NumDirectRefLayers[ nuh_layer_id ] == 1 )  
    NumActiveRefLayerPics = 1  
else  
    NumActiveRefLayerPics = num_inter_layer_ref_pics_minus1 + 1
```

Condition on inter_layer_pred_layer_idc

slice_segment_header() {	Descriptor
....	
if(nuh_layer_id > 0 && NumDirectRefLayers[nuh_layer_id] > 0) {	
inter_layer_pred_enabled_flag	u(1)
if(inter_layer_pred_enabled_flag && NumDirectRefLayers[nuh_layer_id] > 1)	
{	
if(!max_one_active_ref_layer_flag)	
num_inter_layer_ref_pics_minus1	u(v)
if (NumActiveRefLayerPics != NumDirectRefLayers[nuh_layer_id])	
for(i = 0; i < NumActiveRefLayerPics; i++)	
inter_layer_pred_layer_idc[i]	u(v)
}	
}	
....	
}	

- Remove redundant inter_layer_pred_layer_idc[i] when total number of active ref layers is equal to the number of direct dependent layers
- Derivation of inter_layer_pred_layer_idc (F.7.4.3.1.1):

```

if (NumActiveRefLayerPics = NumDirectRefLayers[nuh_layer_id]) {
    for( i = 0; i < NumActiveRefLayerPics; i++ )
        inter_layer_pred_layer_idc[i] = RefLayerId [nuh_layer_id][i]
}

```

Sample/motion prediction syntax elements location

- Propose to move sample/motion prediction syntax elements to PPS extension to save bits
- Keep slice header syntax intact

pps_extension() {	Descriptor
if(nuh_layer_id > 0) {	
NumActiveRefLayerPics = 0	
inter_layer_pred_enabled_flag	u(1)
if(inter_layer_pred_enabled_flag) {	
num_inter_layer_ref_pics_minus1	u(v)
NumActiveRefLayerPics = num_inter_layer_ref_pics_minus1 + 1	
for(i = 0; i < NumActiveRefLayerPics; i++)	
inter_layer_pred_layer_idc[i]	u(v)
}	
if(NumActiveRefLayerPics > 0)	
inter_layer_sample_pred_only_flag	u(1)
if(nuh_layer_id > 0)	
alt_collocated_indication_flag	u(1)
if(alt_collocated_indication_flag)	
collocated_ref_layer_idx	ue(v)
}	
}	

Proposed slice header changes

slice_segment_header() {	Descriptor
...	
if(nuh_layer_id > 0 && NumDirectRefLayers[nuh_layer_id] > 0) {	
inter_layer_pred_enabled_flag	u(1)
if(inter_layer_pred_enabled_flag && NumDirectRefLayers[nuh_layer_id] > 1) {	
if(!max_one_active_ref_layer_flag)	
num_inter_layer_ref_pics_minus1	u(v)
for(i = 0; i < NumActiveRefLayerPics; i++)	
inter_layer_pred_layer_idc[i]	u(v)
}	
}	
if(NumSamplePredRefLayers[nuh_layer_id] > 0 && NumActiveRefLayerPics > 0)	
inter_layer_sample_pred_only_flag	u(1)
...	
if(slice_temporal_mvp_enabled_flag) {	
if(nuh_layer_id > 0 && NumActiveMotionPredRefLayers > 0)	
alt_collocated_indication_flag	u(1)
if(alt_collocated_indication_flag)	
if(NumActiveMotionPredRefLayers > 1)	
collocated_ref_layer_idx	ue(v)
else	
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)
}	
}	
...	
}	

Condition on inter_layer_sample_pred_only_flag

- Inter_layer_sample_pred_only_flag is redundant for IDR pictures

slice_segment_header() {	Descriptor
....	
if(NumSamplePredRefLayers[nuh_layer_id] > 0 && NumActiveRefLayerPics > 0 && (nal_unit_type != IDR_W_RADL && nal_unit_type != IDR_N_LP))	
inter_layer_sample_pred_only_flag	u(1)
....	
}	

inter_layer_sample_pred_only_flag

- TMVP can be bypassed when inter_layer_sample_pred_only_flag is set to 1
- The relative syntax elements are redundant
- Two options proposed to bypass TMVP signals
 1. Relocates inter_layer_sample_pred_only_flag and relative syntax elements before slice_temporal_mvp_enabled_flag
 2. the current signaling order but apply condition on signaling TMVP parameters

Option 1

slice_segment_header() {	Descriptor
...	
if(nuh_layer_id > 0 && NumDirectRefLayers[nuh_layer_id] > 0) {	
inter_layer_pred_enabled_flag	u(1)
if(inter_layer_pred_enabled_flag && NumDirectRefLayers[nuh_layer_id] > 1) {	
if(!max_one_active_ref_layer_flag)	
num_inter_layer_ref_pics_minus1	u(v)
for(i = 0; i < NumActiveRefLayerPics; i++)	
inter_layer_pred_layer_idc[i]	u(v)
}	
if(NumSamplePredRefLayers[nuh_layer_id] > 0 && NumActiveRefLayerPics > 0)	
inter_layer_sample_pred_only_flag	u(1)
if(nal_unit_type != IDR_W_RADL && nal_unit_type != IDR_N_LP) {	
if(sps_temporal_mvp_enabled_flag) (nuh_layer_id > 0 && InterRefEnabledInRPLFlag))	
slice_temporal_mvp_enabled_flag	u(1)
}	
if(nuh_layer_id > 0 && NumDirectRefLayers[nuh_layer_id] > 0) {	
 inter_layer_pred_enabled_flag	u(1)
 if(inter_layer_pred_enabled_flag && NumDirectRefLayers[nuh_layer_id] > 1) {	
 if(!max_one_active_ref_layer_flag)	
 num_inter_layer_ref_pics_minus1	u(v)
 for(i = 0; i < NumActiveRefLayerPics; i++)	
 inter_layer_pred_layer_idc[i]	u(v)
 }	
 if(NumSamplePredRefLayers[nuh_layer_id] > 0 && NumActiveRefLayerPics > 0)	
 inter_layer_sample_pred_only_flag	u(1)
...	
if(slice_temporal_mvp_enabled_flag) {	
if(nuh_layer_id > 0 && NumActiveMotionPredRefLayers > 0)	
alt_collocated_indication_flag	u(1)
if(alt_collocated_indication_flag)	
if(NumActiveMotionPredRefLayers > 1)	
collocated_ref_layer_idx	ue(v)
else {	
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)
}	
...	se(v)
byte_alignment()	
}	

Option 2

slice_segment_header() {	Descriptor
...	
if(slice_temporal_mvp_enabled_flag && InterRefEnabledInRPLFlag) {	
if(nuh_layer_id > 0 && NumActiveMotionPredRefLayers > 0)	
alt_collocated_indication_flag	u(1)
if(alt_collocated_indication_flag)	
if(NumActiveMotionPredRefLayers > 1)	
collocated_ref_layer_idx	ue(v)
else {	
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)
}	
}	
...	
}	

- The derivation process for temporal luma motion vector prediction (SHVC WD2 G.8.5.3.2.7) is modified as follows.
- The variables mvLXCol and availableFlagLXCol are derived as follows:
- If slice_temporal_mvp_enabled_flag is equal to 0 or **InterRefEnabledInRPLFlag is equal to 0**, both components of mvLXCol are set equal to 0 and availableFlagLXCol is set equal to 0.

Conclusion

- Several high level syntax modifications are proposed
- suggest to adopt all proposed items in SHVC and MV-HEVC