



JCTVC-K0206:VPS support for out-of-band signaling and hybrid codec scalability

Jill Boyce, Danny Hong, Wonkap Jang, Stephan Wenger
Vidyo

Ajay Luthra
Motorola Mobility



Introduction

- **VPS syntax changes proposed**

1. Add flag to indicate if scalable/multiview enhancement layer parameters may be sent out-of-band (e.g. in systems layer) rather than in VPS extension
2. Add support for hybrid codec scalability with an AVC base layer

1. Out-of-band via External Means

- **Addition of a `vps_external_means_flag` in the VPS (base specification section, not the VPS extension)**
 - Use a reserved bit
- **Middle box can use this flag to determine whether or not the VPS sufficiently provides the relevant scalability information, for sub-bitstream extraction purposes**

External Means Application Specifications

- **Application specifications or individual profiles may require that the flag have a particular value, for example**
 - Real-time conversational applications may require flag's value shall be 0,
 - Non-real time broadcast applications may allow (or require) that the flag's value shall be 1
 - Systems that need to add or remove the number of layers can do so without parsing into and/or modifying the video layer by setting the flag to 1 and providing relevant information at system layer or by some other means
 - Parsing or modifying video layer can be onerous and/or sometimes not possible as the video layer (s) may be encrypted.

External means proposed VPS syntax & semantics

video_parameter_set_rbsp() {	Descriptor
video_parameter_set_id	u(4)
vps_temporal_id_nesting_flag	u(1)
vps_external_means_flag	u(1)
reserved_zero_12bits	u(12)
vpsmax_num_layers_minus1 //reserved_zero_6bits in the base spec	u(6)
vps_max_sub_layers_minus1	u(3)
profile_level(1, vps_max_sub_layers_minus1)	
next_essential_info_byte_offset //reserved_zero_12bits in the base spec	u(12)

vps_external_means_flag equal to 1 specifies that the ViewId[i], DepthFlag[i], DependencyId[i], QualityId[i], RefLayerId[i][j] variables are set through external means. **vps_external_means_flag** equal to 0 specifies that the ViewId[i], DepthFlag[i], DependencyId[i], QualityId[i], RefLayerId[i][j] variables are set from syntax elements in vps_extension(), as described in sub-section X.X.

vpsmax_num_layers_minus1 plus 1 specifies the maximum number of layers in the coded video sequences referring to the video parameter set described in the vps_extension().

If **vps_external_means_flag** equal to 0, MaxLayers is set equal to vps_num_layers_minus1 + 1. Otherwise the value of MaxLayers is determined through external means.

External Means Proposed VPS Extension Semantics

relative to JCTVC-J1007



- **Introduce variables for the scalability parameters**
 - ViewId, DepthFlag, DependencyId, QualityId
 - Remainder of specification can ignore how values of parameters were determined

For the base layer, ViewId[0], DepthFlag[0], DependencyId[0], QualityId[0] are set equal to 0.

if vps_external_means_flag equal to 0, the following applies

- RefLayerId[i][j] = ref_layer_id[i][j]
- if dimension_type[i][j] equal to 0, ViewId[i] = dimension_id[i][j]
- else if dimension_type[i][j] equal to 1, DepthFlag[i] = dimension_id[i][j]
- else if dimension_type[i][j] equal to 2, DependencyId[i] = dimension_id[i][j]
- else if dimension_type[i][j] equal to 3, QualityId[i] = dimension_id[i][j]

otherwise (vps_external_means_flag equal to 1), the following applies

- ViewId[i], DepthFlag[i], DependencyId[i], QualityId[i], RefLayerId[i][j] are determined through external means for $i > 0$

2. Hybrid Codec Scalability

- **Provides high-level syntax support for HEVC scalability with H.264/AVC base layer**
 - H.264/AVC base layer may conform to any of Annex A, G, H profiles (e.g. includes SVC and MVC)
 - All view, dependency, quality layers of H.264/AVC bitstream are collectively considered to be the HEVC base layer
 - Only base layer (e.g. layer_id = 0) can be coded as H.264/AVC
- **Introduce `avc_base_layer_flag` in VPS extension to indicate the presence of an AVC base layer**
- **Add encapsulation NAL unit type for encapsulated AVC base layer NAL units**
 - Alternatively, systems layer may be used for carriage of AVC base layer bitstream
- **Indicate temporal scalability characteristics of H.264/AVC base layer using HEVC encapsulation NAL unit header's `temporal_id_plus1` syntax element (implies value of TemporalId variable)**

New Encapsulation NAL unit type

- New VCL type NAL unit type value, **AVC_BASE**, added to indicate an encapsulated AVC NAL unit
- H.264/AVC base layer NAL units would be encapsulated in an HEVC NAL unit of this type
- NAL units of this type would contain, in order:
 - HEVC NAL unit header, optional AVC prefix NAL unit, AVC NAL unit

nal_unit_type	Name of nal_unit_type	Content of NAL unit and RBSP syntax structure	NAL unit type class
15	AVC_BASE	Encapsulation of a Rec. ITU-T H.264 ISO/IEC 14496-10 base layer	VCL

When **nal_unit_type** equal to **AVC_BASE**, **layer_id** shall be equal to 0.

Systems Alternative to Encapsulation NUT

- **Alternatively, systems layer may be used for carriage of AVC base layer bitstream**
 - Some system will require the ability to remove the HEVC enhancement layers and forward only the AVC bit streams to the legacy devices without having to extract the AVC stream from the video bit streams
 - In those cases AVC base layer stream and HEVC enhancement stream(s) can be sent a two (or more) separate streams with `vps_external_means_flag` set to 1

Temporal Scalability for Hybrid Codec Scalability

- **Temporal scalability relationship of H.264/AVC base layer and HEVC enhancement layers**
 - When temporal_id is present in AVC base layer (e.g. for profiles defined in Annex G and H), restrict HEVC TemporalId variable to be equal to AVC temporal_id syntax element value
 - When temporal_id not present in AVC base layer (e.g. for profiles defined in Annex A), AVC base layer bitstream complies with TemporalId value of HEVC NAL unit header
 - Use of non-zero TemporalId values indicates AVC base layer bitstream's use of reference picture lists enables operation of AVC's sub-bitstream extraction process
- **HEVC currently restricts all NAL units of same access unit to have same value of TemporalId. Propose to loosen that restriction for encapsulated AVC base layer NAL units.**
- **Middle box may safely perform sub-bitstream extraction using HEVC NAL unit header values (temporal_id_plus1, layer_id)**

Add avc_base_layer_flag to VPS Extension

vps_extension() {	Descriptor
while(!byte_aligned())	
vps_extension_byte_alignment_reserved_zero_bit	u(1)
avc_base_layer_flag	u(1)
// layer specific information	

avc_base_layer_flag equal to 1 specifies that the base layer conforms to Rec. ITU-T H.264 | ISO/IEC 14496-10, equal to 0 specifies that it conforms to this specification. When temporal_id present in the Rec. ITU-T H.264 | ISO/IEC 14496-10 conforming base layer, temporal_id shall be equal to TemporalId. In the Rec. ITU-T H.264 | ISO/IEC 14496-10 conforming base layer, after applying the Rec. ITU-T H.264 | ISO/IEC 14496-10 decoding process for reference picture lists construction the output reference picture lists refPicList0 and refPicList1 (when applicable) shall not contain any pictures for which the TemporalId is greater than TemporalId of the current picture. All sub-bitstreams of the Rec. ITU-T H.264 | ISO/IEC 14496-10 conforming base layer, that can be derived using the sub-bitstream extraction process as specified in Rec. ITUT H.264 | ISO/IEC 14496-10 subclause G.8.8.1 with any value for temporal_id as the input shall result in a set of coded video sequences, with each coded video sequence conforming to one or more of the profiles specified in Rec. ITUT H.264 | ISO/IEC 14496-10 Annexes A, G and H.

Conclusion

- **Recommend to adopt proposed flags and semantics to enable**
 - Out-of-band signaling of enhancement layer parameters
 - Hybrid codec scalability with H.264/AVC base layer