JCTVC-F290: Scalability Support in HEVC

Danny Hong, Wonkap Jang, Jill Boyce, Adeel Abbas

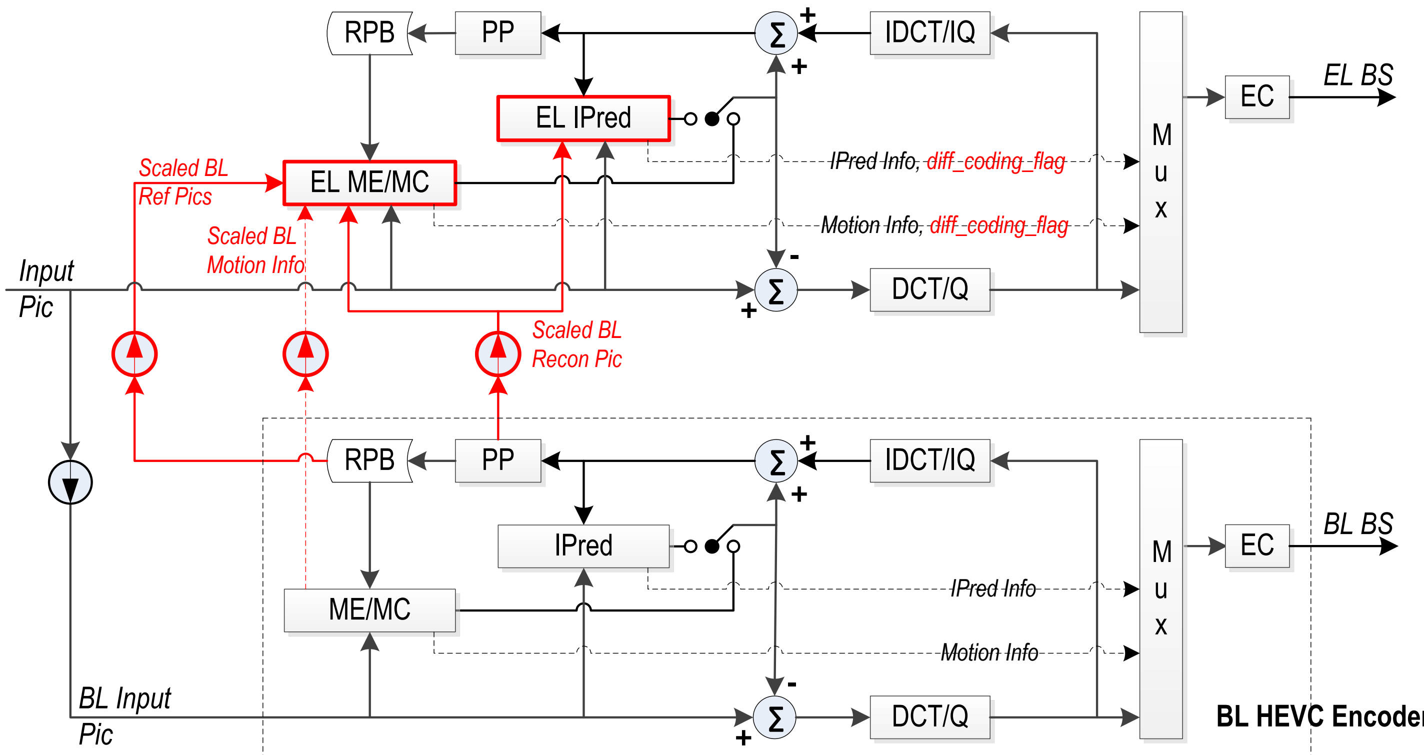
Vidyo

Introduction

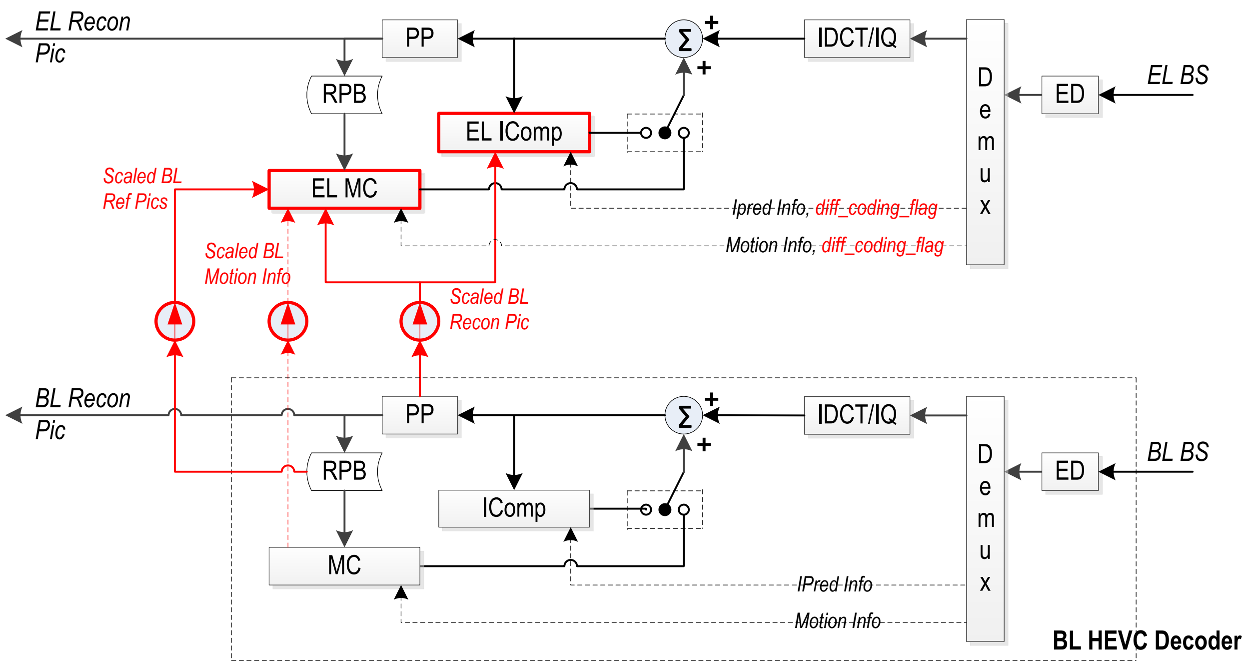
- **A scalable coding architecture is proposed**
  - Allows spatial scalability at any resolution ratio (or CGS with no resolution change)
  - Experimental results for intra and inter, 2 spatial layers, 1:2 resolution ratio per dimension
  - Design goal to minimize changes to HEVC codec
  - Architecture could be applied to an H.264/AVC backwards-compatible base layer

Block diagram – 2 layer encoder

- **Any downsampling method may be used**
  - JSVM method used in experiments
- **DCTIF filter used for upsampling**



Block diagram – 2 layer decoder



- **Also extensible to 3 or more layers**

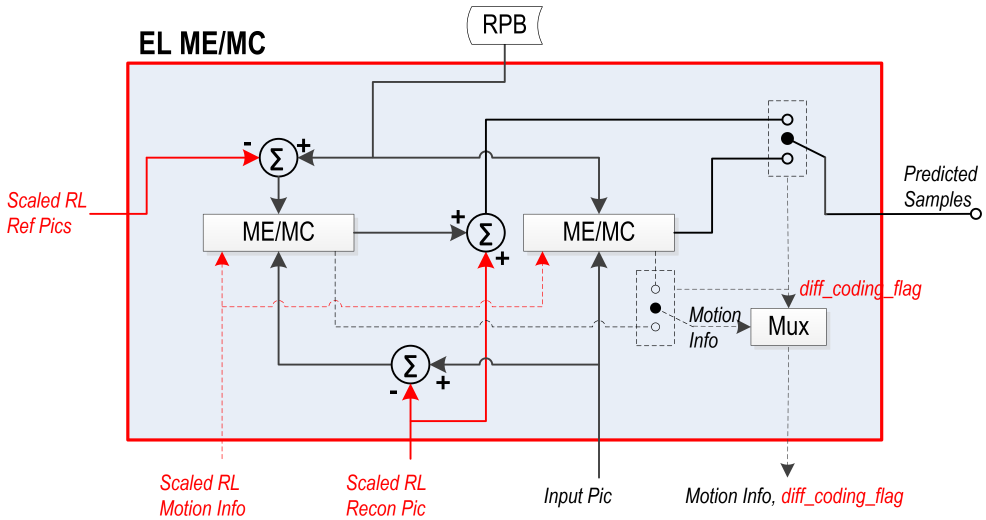
Two main coding modes

- **Difference coding mode**
  - Base layer coded pixels upsampled and subtracted from high resolution input, forming difference values
  - Difference values are coded using normal HEVC
- **Pixel coding mode**
  - High resolution input pixels coded in non-scalable manner

- **Signaling at slice level or CU level, using diff_coding_flag**
- **Double loop decoding required**
  - SVC only requires double loop decoding for intra
  - Even with SVC, some applications benefit from broader use of double loop decoding

Motion estimation/compensation

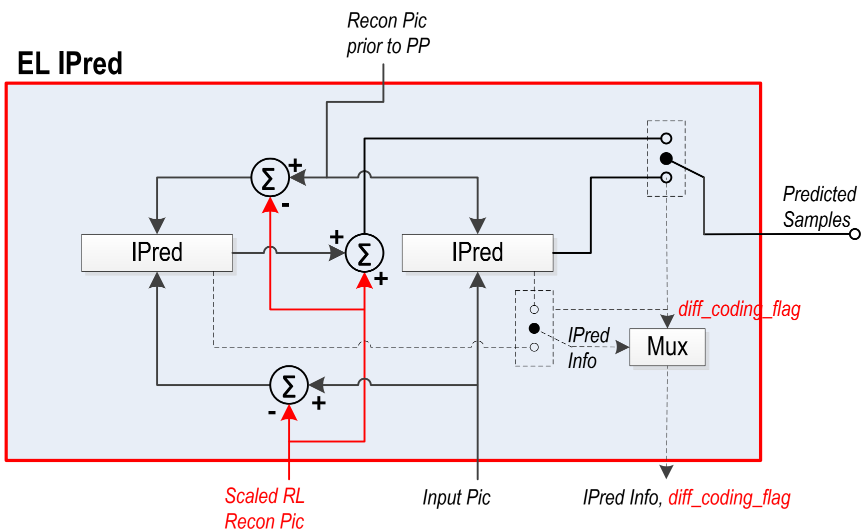
- **When difference coding mode used, difference values (high res input picture minus upsampled coded reference layer picture) predicted from difference values of previously coded picture**



- **Reference block(s) in reference picture may have been coded in either pixel coding mode or difference coding mode**
  - Reference picture difference values easily derived

Intra prediction

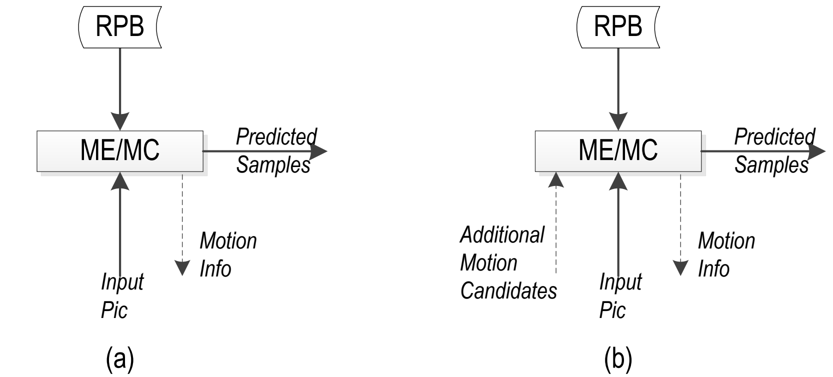
- **When difference coding mode used, difference values spatially predicted from neighboring difference blocks**



- **Spatial neighboring blocks may have been coded in either pixel coding mode or difference coding mode**
  - Difference values easily derived

Inter-layer motion prediction

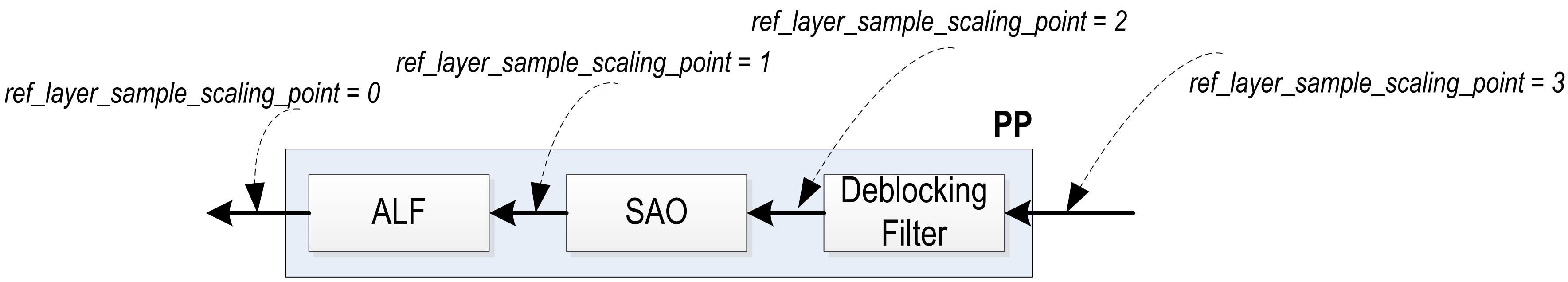
- **Base layer Motion Vector added to the motion prediction list, at the end, if available**
- **Center of the current block used to find co-located reference layer MV**
- **Scale MV by resolution scaling factor**
- **No new syntax in enhancement layer**



(a) Current ME/MC (b) Proposed ME/MC

Reference layer sample point

- **Choose whether upsampling of the reference layer done before or after the in-loop filtering modules**
  - Analogous to disable_inter_layer_deblocking_filter_idc in SVC



Syntax changes

- **Dependency Parameter Set (initially proposed in JCTVC-E279)**
  - Use dependency_id instead of spatial_id, similar to SVC, because of similar handling of spatial and CGS layers
- **Picture parameter set**
  - Add ref_layer_info_present_flag, ref_layer_sample_scaling_point
- **Slice header**
  - Add slice_skip_flag, adaptive_diff_coding_flag
- **CU**
  - Add diff_coding_flag
- **No sub-CU syntax changes**

H.264/AVC base layer mixed codec option

- **Coding of enhancement layer less tightly coupled to base layer coding than in SVC**
- **H.264/AVC base layer may be supported using same architecture, with only high level syntax changes to enhancement layer**
  - Option 1
    - Use different HEVC NAL unit types than H.264/AVC
      - Similar to proposal in JCTVC-E375
    - Add codec type field to Dependency Parameter Set
    - Only base layer can be of a different codec type, and only H.264/AVC supported
      - May be SVC (Annex G) with multiple dependency or quality layers of its own
  - Option 2
    - Encapsulation NAL unit type defined
    - Encapsulation NALs would include sequences compliant with a non-HEVC standard
    - Allows for any coding standard
    - Allows for multiple scalable or simulcast layers

Experimental conditions

- **2 spatial layers, 1:2 ratio per dimension**
- **HM 3.0 modified**
- **Common test conditions used**
- **Base layer QP set to 2 below enhancement layer QP**
- **Scalability compared to simulcast and single layer**
- **Comparison to JSVM version 9.18**
  - CABAC for both layers for HE comparison
  - CAVLC for both layers for LoCo comparison

- **See JCTVC-F292 for a discussion of metrics for assessing scalable video coding efficiency**
  - Curve fitting of simulcast curve, to match scalable enhan layer PSNR

Scalable HEVC experimental results

Vidyo™ Personal Telepresence

Simulcast vs Scalable

	Intra			Intra LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	-25.6	-25.4	-25.5	-24.4	-24.2	-24.2
Class B	-21.3	-21.3	-21.3	-19.2	-19.0	-19.0
Class C	-13.3	-13.2	-13.2	-11.6	-11.6	-11.6
Class D	-11.4	-11.3	-11.4	-10.2	-10.2	-10.2
Class E	-22.0	-21.9	-21.9	-20.1	-20.1	-20.1
All	-17.9	-17.9	-17.9	-16.3	-16.1	-16.2

	Random access			Random access LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	-21.9	-21.8	-21.8	-21.4	-21.3	-21.3
Class B	-20.1	-20.2	-20.1	-19.4	-19.4	-19.2
Class C	-14.9	-14.9	-14.9	-14.2	-14.2	-14.2
Class D	-9.5	-9.5	-9.5	-9.0	-9.1	-9.1
Class E						
All	-16.1	-16.2	-16.1	-15.5	-15.5	-15.5

	Low delay			Low delay LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A						
Class B	-15.9	-16.0	-16.1	-15.8	-15.8	-15.8
Class C	-12.2	-12.2	-12.2	-11.6	-11.6	-11.6
Class D	-7.4	-7.4	-7.4	-7.1	-7.1	-7.1
Class E	-7.9	-7.8	-7.9	-8.1	-8.1	-8.2
All	-11.4	-11.4	-11.4	-11.1	-11.1	-11.1

Single-Layer vs Scalable

	Intra			Intra LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	16.2	15.7	9.8	17.6	19.6	15.9
Class B	14.5	18.8	15.9	18.1	20.9	18.5
Class C	21.1	26.8	28.9	23.2	25.3	26.1
Class D	22.0	29.1	30.9	23.4	26.0	26.4
Class E	22.1	15.7	20.7	24.4	17.8	20.3
All	19.1	22.0	22.2	21.4	22.3	22.0

	Random access			Random access LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	22.5	27.1	20.9	22.7	33.0	28.0
Class B	15.8	22.9	20.7	16.8	25.4	23.8
Class C	20.7	28.1	30.4	21.5	28.6	29.4
Class D	24.3	32.2	32.7	24.9	31.7	31.8
Class E						
All	20.3	27.3	26.5	21.0	29.0	28.0

	Low delay			Low delay LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A						
Class B	21.3	22.6	23.0	21.6	26.2	28.8
Class C	24.3	28.6	29.5	24.8	31.5	32.2
Class D	26.5	30.9	32.0	26.8	32.8	33.1
Class E	38.4	39.3	40.3	37.3	41.2	44.2
All	26.6	29.3	30.1	26.7	32.0	33.6

JSVM results

using HEVC test conditions, for comparison

Vidyo™ Personal Telepresence

Simulcast vs Scalable

	Intra			Intra LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	-31.3	-30.6	-30.8	-29.7	-28.9	-29.2
Class B	-25.8	-25.3	-25.4	-23.7	-23.1	-23.2
Class C	-16.9	-16.5	-16.5	-14.7	-14.2	-14.2
Class D	-15.1	-14.6	-14.7	-12.6	-12.1	-12.2
Class E	-25.5	-25.5	-25.3	-24.8	-24.7	-24.5
All	-22.0	-21.6	-21.6	-20.1	-19.6	-19.7

	Random access			Random access LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	-19.9	-19.3	-19.4	-18.7	-18.1	-18.3
Class B	-19.3	-19.0	-18.9	-17.8	-17.6	-17.5
Class C	-13.4	-13.1	-13.1	-12.4	-12.1	-12.1
Class D	-8.9	-8.7	-8.7	-7.9	-7.7	-7.7
Class E						
All	-15.0	-14.7	-14.7	-13.8	-13.6	-13.6

	Low delay			Low delay LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A						
Class B	-15.8	-15.5	-15.5	-14.6	-14.4	-14.4
Class C	-11.1	-10.9	-10.9	-10.3	-10.1	-10.1
Class D	-7.1	-7.0	-7.0	-6.4	-6.3	-6.3
Class E	-10.3	-10.3	-10.3	-8.6	-8.6	-8.6
All	-11.4	-11.2	-11.2	-10.4	-10.2	-10.2

Single-Layer vs Scalable

	Intra			Intra LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	2.5	4.6	3.8	3.4	5.5	4.7
Class B	4.5	5.4	4.4	6.7	7.5	6.5
Class C	14.0	15.2	15.9	16.6	17.8	18.4
Class D	15.7	16.6	17.7	18.7	19.5	20.6
Class E	12.1	6.2	5.5	11.1	5.7	5.1
All	10.1	10.1	10.0	11.9	11.9	11.8

	Random access			Random access LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	22.1	18.5	17.0	23.1	19.4	18.1
Class B	14.8	11.1	9.2	16.3	12.7	10.8
Class C	20.1	19.0	19.2	21.2	20.1	20.3
Class D	24.6	23.4	24.2	25.8	24.6	25.3
Class E						
All	19.8	17.5	16.9	21.1	18.8	18.2

	Low delay			Low delay LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A						
Class B	18.7	14.6	13.9	20.0	15.9	15.7
Class C	22.6	20.3	20.4	23.5	21.3	21.4
Class D	26.3	24.0	24.2	27.1	25.1	25.1
Class E	30.5	20.2	18.7	32.5	22.3	21.0
All	23.8	19.4	19.0	25.0	20.8	20.5

Experimental results for HEVC vs. SVC

Vidyo™ Personal Telepresence

- **Proposed scalable HEVC codec has similar bitrate savings vs. simulcast as SVC for Random Access and Low Delay, for tested conditions**
  - Relatively more savings for high resolutions than for lower resolutions
- **Proposed scalable HEVC codec has lower bitrate savings percentages vs. simulcast as SVC for Intra**

Mixed codec experimental results

- **Experiments performed with same H.264/AVC base layer as used in the SVC JSVM experiments, but with an HEVC enhancement layer**
  - As expected, architecture with HEVC enhancement layer more efficient than SVC enhancement layer

	Intra			Intra LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	#VALUE!	#VALUE!	#VALUE!	-15.6	-12.1	-12.2
Class B	#VALUE!	#VALUE!	#VALUE!	-17.3	-16.5	-17.0
Class C	#VALUE!	#VALUE!	#VALUE!	-15.6	-13.4	-13.9
Class D	#VALUE!	#VALUE!	#VALUE!	-13.0	-8.4	-8.7
Class E	#VALUE!	#VALUE!	#VALUE!	-19.6	-22.4	-23.5
All	#VALUE!	#VALUE!	#VALUE!	-16.2	-14.5	-15.0

	Low delay			Low delay LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A						
Class B	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
Class C	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
Class D	#VALUE!	#VALUE!	#VALUE!	-29.7	-29.7	-29.7
Class E	#VALUE!	#VALUE!	#VALUE!	-35.1	-35.1	-35.1
All	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!

	Random access			Random access LoCo		
	Y BD-rate	U BD-rate	V BD-rate	Y BD-rate	U BD-rate	V BD-rate
Class A	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
Class B	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
Class C	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
Class D	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
Class E						
All	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!

Conclusions

- **Scalable coding extension to HEVC proposed**
- **Similar BD-bitrate savings vs. simulcast for RA and LD test cases as JSVM**
- **Flexible design**
  - Design goal of few changes vs. single layer HEVC design
  - Any scaling factor can be supported
  - Can be used with H.264/AVC base layer