

REDEFINING MOBILITY



JCTVC-K0035/K0036 Description of scalable video coding technology proposal by Qualcomm

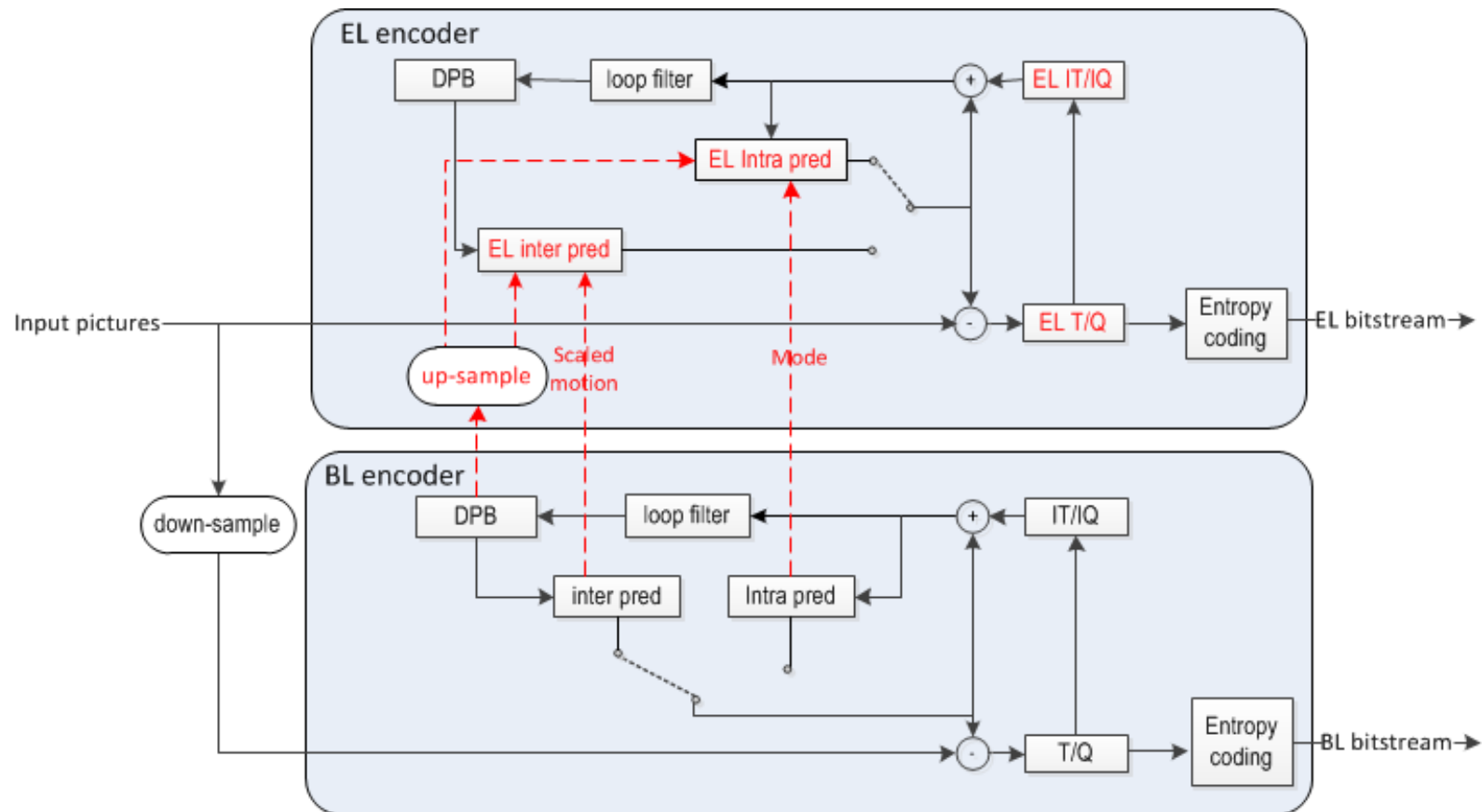
Jianle Chen, Krishna Rapaka, Xiang Li, Vadim Seregin, Liwei Guo, Marta Karczewicz, Geert Van der Auwera, Joel Sole, Xianglin Wang, Chengjie Tu, Ying Chen

Overview

- Introduction
- Algorithm Description
 - Inter-layer intra prediction
 - Inter-layer inter prediction
 - Up-sampling filter
 - Inter-layer syntax prediction
 - Spatial transforms/Entropy coding
- Compression Performance
- Complexity Analysis
- Summary

Introduction

- Multi-loop decoding framework
 - All information of base layer (BL) is available for enhancement layer (EL) coding



Intra Prediction (1)

- Intra prediction

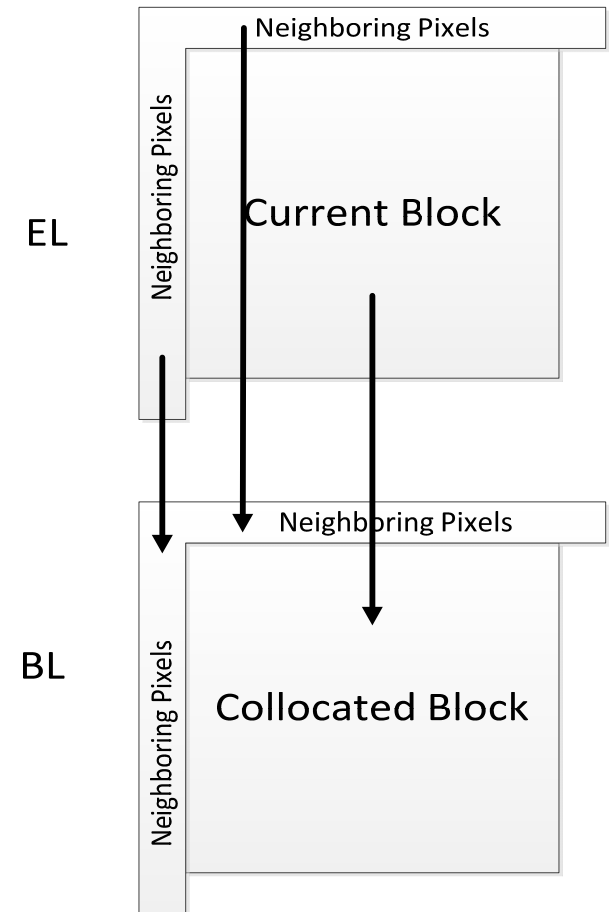
- Prediction process is same as in HEVC (35 modes used) except DC mode
 - DC prediction - average of the collocated BL block
- Use intra prediction mode of the collocated BL block (colBaseDir) to refine Most Probable Intra Prediction Mode (MPM) list
 - When colBaseDir is not a DC mode
 - And when colBaseDir is not in the existing MPM list

```
candModeList[0] = colBaseDir  
candModeList[1] = candModeList[0]  
candModeList[2] = candModeList[1]
```

- Intra-BL prediction - pixels in collocated BL block used as prediction

Intra Prediction (2)

- Intra residual prediction (difference domain prediction)
 - Collocated BL block signal and difference domain prediction are added
 - Difference domain prediction
 - The reference pixels in difference domain are obtained by subtracting collocated BL pixels from the EL reference pixels
 - Intra prediction process in difference domain is the same as in HEVC except planar mode
 - » Prediction of right-bottom portion of the PU is set to zero
 - MPM List.
 - The horizontal and vertical modes are given higher priority than the DC mode during MPM List derivation



Generalized Inter Residual Prediction (GRP)

- The prediction P of GRP mode is:

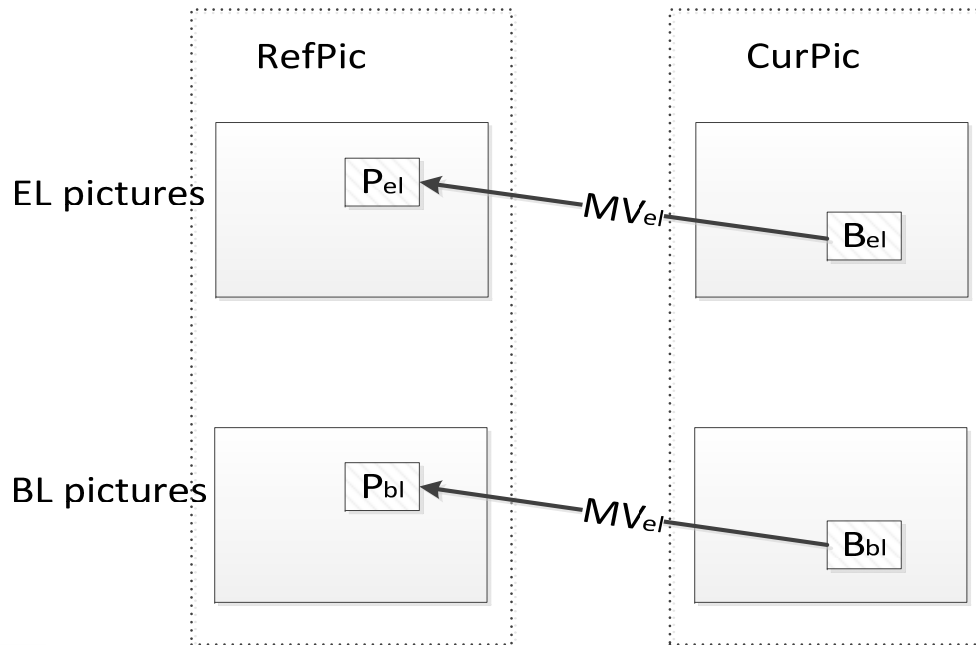
$$P = P_{el} + wR_{bl} = P_{el} + w \cdot (B_{bl} - P_{bl}), \quad w = 0, \frac{1}{2}, 1$$

B_{el} : Current block

B_{bl} : Collocated BL block

P_{el} : EL inter prediction with EL motion vector

P_{bl} : BL inter prediction with EL motion vector

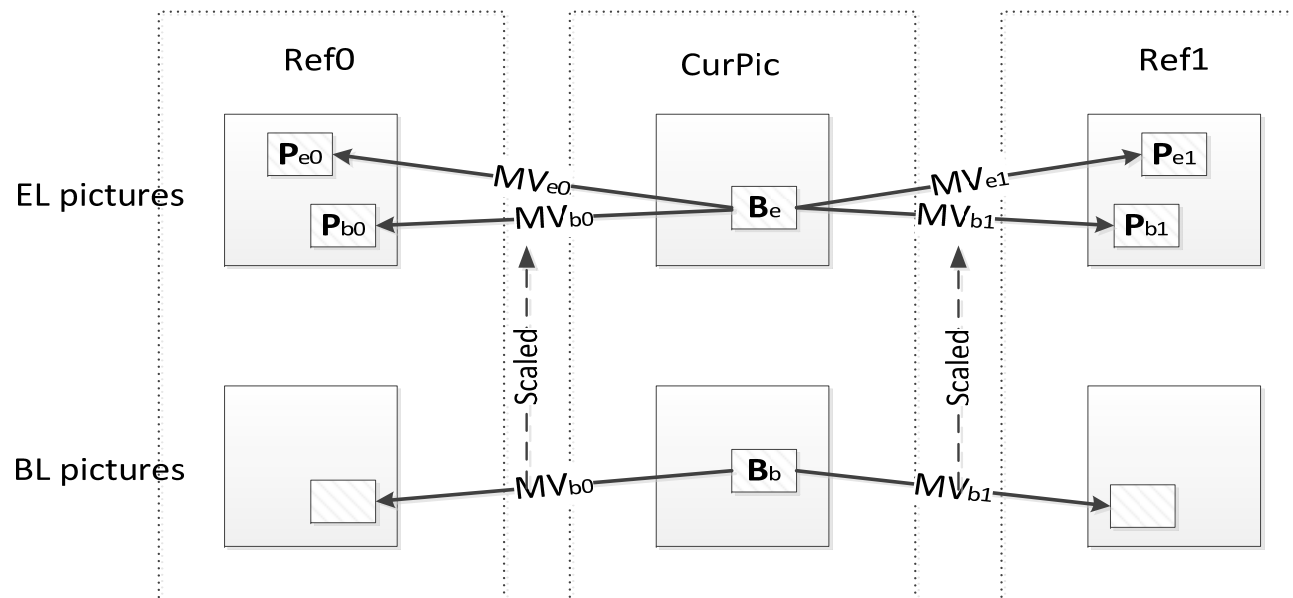


Multi-Hypothesis Motion Prediction

Note: only used in K0036

- Up to two additional prediction hypotheses in the EL reference pictures are obtained by using collocated BL motion vectors
 - PU level *multi_hypothesis_flag* indicates the usage of this mode

$$P = (((P_{e0} + P_{b0} + 1)/2) + ((P_{e1} + P_{b1} + 1)/2) + 1)/2$$



Combined Prediction

Note: only used in K0036

- Combine intra-BL prediction and EL intra/inter prediction

$$P = (P_{el} + P_{bl} + 1) / 2$$

- Intra prediction
 - Only the first mode in MPM list used
- Inter prediction
 - Only applied when GRP weight $w = 1$
- Only applied to luma component
 - EL intra/inter prediction is used for chroma component

Inferred Base Layer Mode

Note: only used in K0036

- When *baselayer_mode_flag* is true, prediction information for the current CU is inferred from its collocated base layer blocks
 - For each 4x4 block, prediction information is copied from its collocated BL block
 - When collocated block is inter coded
 - » Motion information is scaled and copied
 - » Prediction is generated with the scaled MV
 - When collocated block is intra coded
 - » Prediction is generated as Intra BL mode

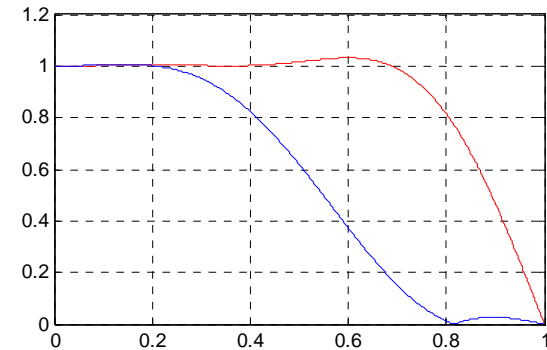
Up-Sampling Filter

■ Luma filter

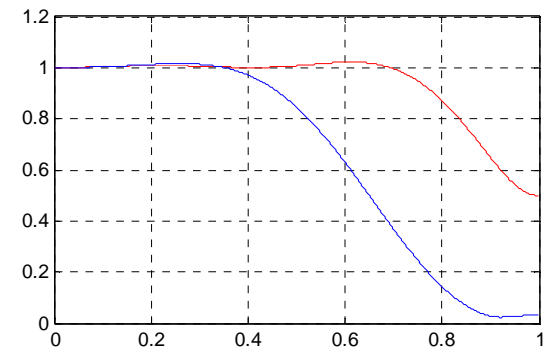
- Coefficients are 6 bits accuracy
- Intra (residual) up-sampling filter
 - Fixed 8-tap 2D separable filter
- Inter (residual) up-sampling filter
 - Fixed 8-tap 2D separable filter
- Intra-BL prediction
 - Adaptive 8-tap 2D separable filter
 - Coefficients are signaled once a frame
- Filter is also applied to integer position
 - Filtering process applied to both SNR and spatial scalability cases

■ Chroma filter

- 4-tap 2D separable filter

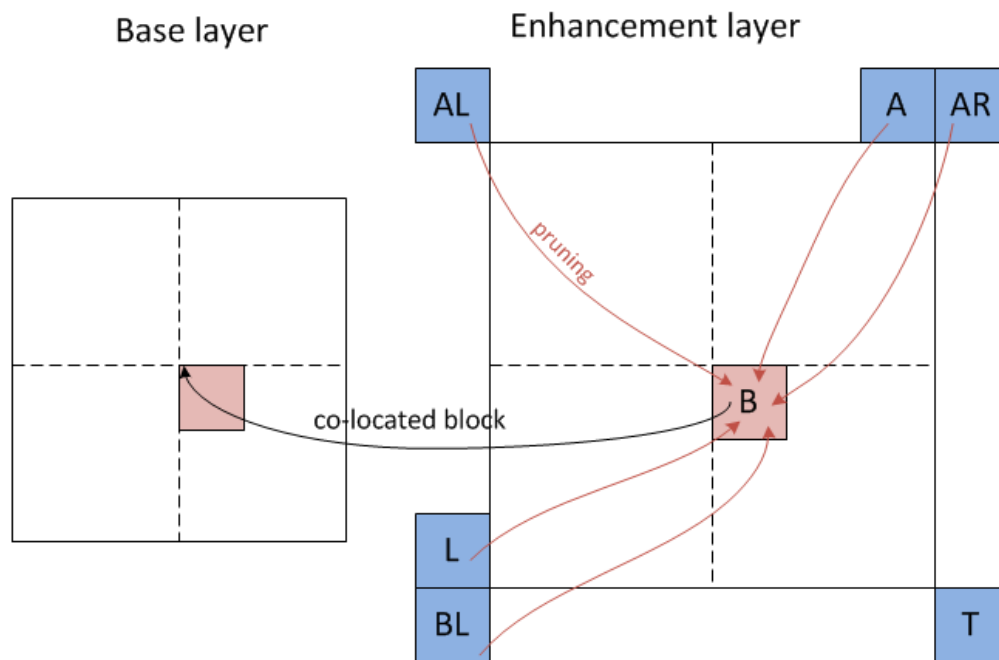


Filter phase = 8



Filter phase = 5

Inter-Layer Motion Parameter Prediction



Merge list (total 5)	AMVP list (total 2)
Base layer MV	Left
Left	Above
Above	TMVP
Above right	Base layer MV
Left bottom (if candNum<4)	Zero candidates
Above left (if candNum<4)	
TMVP	
Combined candidates	
Offset MV	
Zero candidates	

■ Merge

- BL MV is inserted as the first candidate, pruning is done for each spatial candidates with co-located BL MV
- HEVC pruning keeps unchanged

■ AMVP

- BL MV is inserted after TMVP

Spatial Transform

Note: 5 transforms used in K0036 and 3 transforms used in K0035

- Inter-layer predicted residue has different characteristics to the HEVC residue
 - Intra-BL mode: additional 2 DCTs and 2 DSTs used
 - Transform index is coded at CU level with truncated unary codes

index	Transform
0	DST Type-3
1	DCT Type-2
2	DCT Type-8
3	DST Type-4
4	DCT Type-3

- Inter residual prediction mode
 - DST Type-3 is applied instead of DCT Type-2
 - » GRP weight $w = 1$
 - » CU size $\leq 16 \times 16$

Miscellaneous

- Inter layer CU split prediction
 - The current CU is not further split when its size is smaller than the size of collocated BL CU and spatial neighboring CUs
- Entropy coding
 - Remove single context assigning to the high-frequency coefficient for large TU
 - Additional coefficients scan order derivation process for 4x4 and 8x8 TU
 - Intra mode dependent coefficient scan
 - The horizontal gradient G_h and vertical gradient G_v of collocated BL block
 - » horizontal scan if $G_h > 2G_v$
 - » vertical scan if $G_v > 2G_h$
- De-blocking filter
 - De-blocking process is the same as HEVC except setting BS for the new modes

Encoding setting

- The same GOP structure in enhancement layer and base layer
- Lambda for enhancement layer
 - λ is determined not only by enhancement layer QP but also base layer picture quality
 - Base layer QP
 - Spatial resolution ratio
- QP hierarchy for enhancement layer (based on JCTVC-J0242)
$$QP' = \min(QP_{anchor} + 2, \max(QP_{anchor}, 4.2005 \cdot \ln(\lambda) + 1.7122 + 0.5))$$
- JCTCV-K0036
 - Enable AMP in EL
 - Enable HM6.1- ALF in EL

Test configurations

- ☒ Category 1 (HEVC base layer) spatial scalability
- ☒ Category 1 (HEVC base layer) intra-only spatial scalability
- ☒ Category 1 (HEVC base layer) SNR scalability
- ☐ Category 2 (AVC base layer) spatial scalability

- ☒ The proposal obeys the constraints under section 5 of the CfP (if box is not ticked, explain cases where constraints are violated)

Experiment Results

- The actual EL rate is around 0.5% lower than the target bit rate

Test configuration	<i>Total(EL+BL) rate</i>			<i>EL rate</i>		
	BD-rate saving					
	Y	Cb	Cr	Y	Cb	Cr
Intra 2x	-26.5%	-22.9%	-23.4%	-37.6%	-33.7%	-34.2%
Intra 1.5x	-34.1%	-30.8%	-31.3%	-56.1%	-52.5%	-53.0%
RA 2x	-25.6%	-18.9%	-19.2%	-35.8%	-28.3%	-28.6%
RA 1.5x	-33.4%	-27.2%	-27.0%	-52.6%	-46.2%	-46.0%
RA SNR	-30.6%	-22.5%	-22.2%	-44.0%	-34.8%	-34.4%
RA Average	-29.8%	-22.9%	-22.8%	-44.2%	-36.5%	-36.3%

Encoding complexity

- Encoding time (geometric mean)

- EL and BL encoding time of proposed SVC encoder
- EL encoding time of HM6.1 encoder

Test configuration	HM61 (hours)	Proposed (hours)	Ratio
Intra 2x	2.16	9.44	4.36
Intra1.5x	1.92	8.14	4.25
RA 2x	7.00	18.53	2.65
RA 1.5x	6.24	16.36	2.62
RA SNR	7.36	25.47	3.46

- Coding mode selection at encoder side

- ❖ Intra prediction

- Intra-BL
 - Combined prediction
 - Intra directional mode:
 - *intra_resi_pred_flag* = 0
 - *intra_resi_pred_flag* = 1

- ❖ Inter prediction

- Motion estimation with GRP weight $w = 0$ same as HM6.1
 - RD check
 - Inter-skip & Inter Mode
 - GRP weights 0, 1/2, 1
 - Multi-hypothesis
 - Combined prediction
 - Inferred base layer mode

Decoding complexity

■ Decoding time

- YUV output for EL was enabled and reference input was disabled
- The geometric mean
 - EL and BL decoding time of proposed SVC decoder
 - EL decoding time of HM6.1 decoder
- A single Windows PC with Intel Xeon W3530 4 cores CPU (2.8GHz), 6GB memory and SATA hard disk with 7200 RPM

Test configuration	HM61 (seconds)	Proposed (seconds)	Ratio
Intra 2x	80.73	167.08	2.07
Intra1.5x	69.74	152.93	2.19
RA 2x	48.98	110.60	2.26
RA 1.5x	43.19	112.25	2.60
RA SNR	52.39	142.25	2.72

Tools setting in JCTVC-K0035

- No AMP
- No ALF
- Disable 'inferred base layer mode'
- Disable 'multi-hypothesis motion prediction'
- Disable 'combined prediction'
- Three transforms is applied for intra-BL mode CU

Experiment Results (JCTVC-K0035)

- The actual EL rate is around 0.5% lower than the target bit rate

Test configuration	<i>Total(EL+BL) rate</i>			<i>EL rate</i>		
	BD-rate saving					
	Y	Cb	Cr	Y	Cb	Cr
Intra 2x	-25.6%	-22.8%	-23.3%	-36.7%	-33.7%	-34.1%
Intra 1.5x	-33.5%	-30.9%	-31.2%	-55.5%	-52.6%	-52.9%
RA 2x	-23.0%	-15.7%	-16.4%	-33.0%	-24.6%	-25.4%
RA 1.5x	-31.1%	-23.9%	-24.1%	-50.2%	-42.5%	-42.7%
RA SNR	-28.2%	-18.6%	-18.3%	-41.4%	-30.3%	-29.8%
RA Average	-27.5%	-19.4%	-19.6%	-41.5%	-32.5%	-32.6%

Encoding complexity (JCTVC-K0035)

- Encoding time (geometric mean)

- EL and BL encoding time of proposed SVC encoder
- EL encoding time of HM6.1 encoder

Test configuration	HM61 (hours)	Proposed (hours)	Ratio
Intra 2x	2.16	7.30	3.37
Intra1.5x	1.92	6.38	3.33
RA 2x	7.00	11.58	1.65
RA 1.5x	6.24	10.93	1.75
RA SNR	7.36	17.39	2.36

- Complexity analysis for coding mode selection

- ❖ Intra prediction

- Intra-BL
- Intra direction mode:
 - *intra_resi_pred_flag* = 0
 - *intra_resi_pred_flag* = 1

- ❖ Inter prediction

- Motion estimation with GRP weight
 $w = 0$, same as HEVC
- RD check
 - Inter-skip & Inter Mode
 - GRP weights 0, 1/2, 1

Decoding complexity (JCTVC-K0035)

■ Decoding time

- YUV output for EL was enabled and reference input was disabled
- The geometric mean
 - EL and BL decoding time of proposed SVC decoder
 - EL decoding time of HM6.1 decoder
- A single Windows PC with Intel Xeon W3530 4 cores CPU (2.8GHz), 6GB memory and SATA hard disk with 7200 RPM

Test configuration	HM61 (seconds)	Proposed (seconds)	Ratio
Intra 2x	80.73	155.55	1.93
Intra1.5x	69.74	143.88	2.06
RA 2x	48.98	99.18	2.02
RA 1.5x	43.19	100.24	2.32
RA SNR	52.39	127.94	2.44

Additional tests based on JCTVC-K0035

- Enable AMP + ALF + IBDI

Test configuration	Total(EL+BL) rate (10 bits)			Total(EL+BL) rate (8 bits)		
	BD-rate saving					
	Y	Cb	Cr	Y	Cb	Cr
Intra 2x	-26.8%	-24.1%	-24.8%	-26.0%	-22.3%	-22.7%
Intra 1.5x	-34.3%	-32.2%	-32.8%	-33.6%	-30.7%	-31.1%
RA 2x	-26.7%	-23.3%	-24.5%	-25.8%	-19.8%	-19.9%
RA 1.5x	-34.6%	-31.9%	-32.1%	-33.7%	-28.5%	-27.9%
RA SNR	-31.5%	-27.4%	-28.0%	-30.4%	-23.5%	-22.7%
RA Average	-31.0%	-27.5%	-28.2%	-30.0%	-23.9%	-23.5%

Encoding time

Test configuration	HM61 (hours)	Proposed (hours)	Ratio
Intra 2x	2.16	7.43	3.43
Intra1.5x	1.92	6.44	3.36
RA 2x	7.00	13.08	1.87
RA 1.5x	6.24	12.10	1.94
RA SNR	7.36	19.09	2.60

Summary

■ JCTVC-K0035:

- Intra
- Intra-BL
- Intra residual prediction
- Inter residual prediction
- Inter-layer MV prediction
- Additional 2 transforms

■ Additional test of JCTVC-K0035:

- AMP
- ALF
- IBDI

■ JCTVC-K0036:

- Intra
- Intra-BL
- Intra residual prediction
- Inter residual prediction
- Multiple-hypothesis
- Combined prediction
- Inferred base layer mode
- Inter-layer MV prediction
- Additional 4 transforms
- AMP
- ALF



■ Thanks