

Description of scalable video coding technology proposal by ETRI and Kwangwoon Univ. (JCTVC-K0037)

Jung Won Kang, Hahyun Lee, Jinho Lee (ETRI)

Junghak Nam, Hyomin Choi, Donggyu Sim (Kwangwoon Univ.)

Contents

- ❑ **Overview**
- ❑ **Algorithm Description**
 - ❖ Inter-layer texture prediction
 - ❖ Inter-layer motion vector prediction
 - ❖ Up-sampling filters
- ❑ **Compression Performance**
- ❑ **Complexity Analysis**
- ❑ **Summary**

Overview

❑ The proposal based on multiple loop decoding design

- ❖ Be easily extended to support multi-view scalability at the same time
 - since multiple scalable structure which perfectly reconstruct frame of all layer can display any views of the multi-view configuration as view scalability
- ❖ Can support coding standard scalability (AVC/H.264 base layer)

❑ Categories the proposal covers

- ❖ Category 1 (HEVC base layer) spatial scalability
- ❖ Category 1 (HEVC base layer) intra-only spatial scalability
- ❖ Category 1 (HEVC base layer) SNR scalability

Overview

❑ Features applied to the proposal

❖ Inter-layer texture prediction

- A reconstructed picture of the reference layer is used as the reference picture for the corresponding picture of the enhancement layer

❖ Merge/skip & motion vector prediction candidates

- The construction process and positions of merge/skip and motion vector prediction candidates are modified to consider the corresponding PU on the reference layer in ME/MC of the enhancement layer

❖ ALF is applied to the enhancement layer

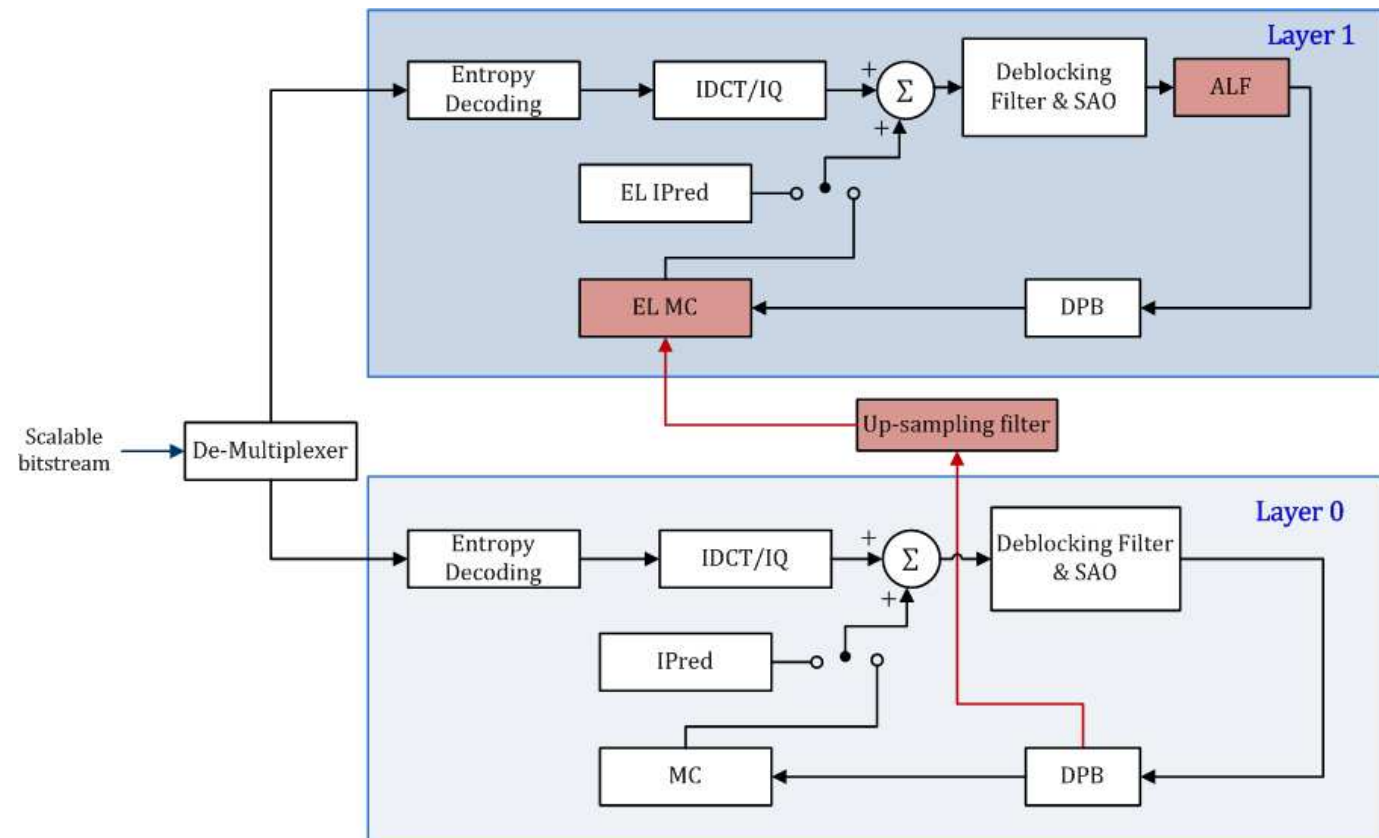
❖ Up-sampling filters

- Luma: 8 taps DCT-IF filters with 6 bits filter coefficients representation
- Chroma: 4 taps DCT-IF filters with 6 bits filter coefficients representation

Overview

❑ Decoder Block Diagram

- ❖ Each layer employs HEVC tools
- ❖ Inter-layer predictions for the enhancement layer
- ❖ Support spatial, SNR and temporal scalabilities



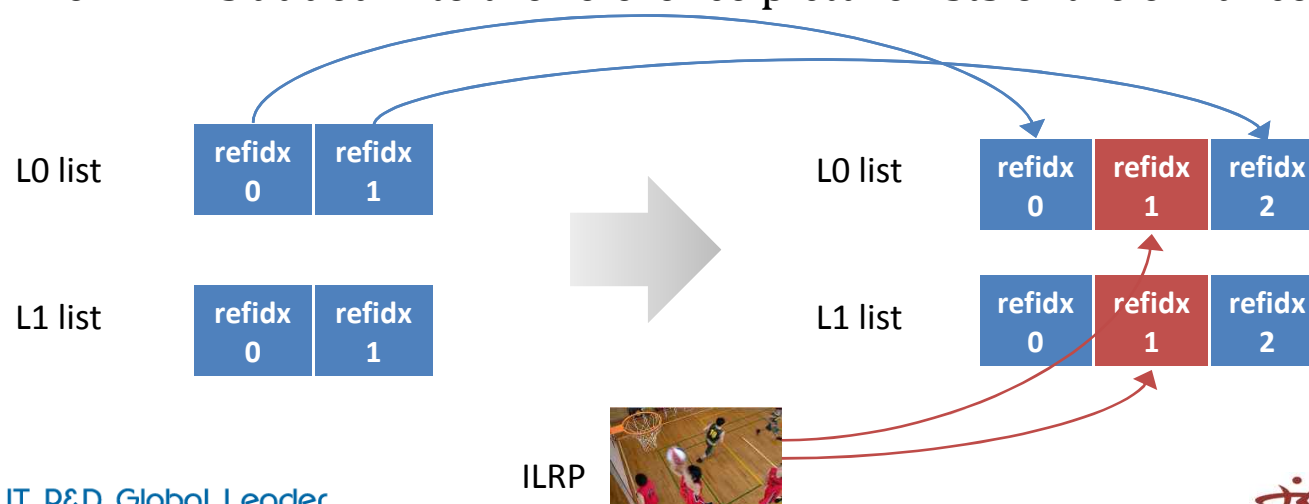
Inter-layer texture prediction

❑ ILRP (Inter-layer reference picture)

- ❖ Reconstructed (upsampled) picture of the reference layer
- ❖ ILRP is added into the reference picture lists for the corresponding picture of the enhancement layer

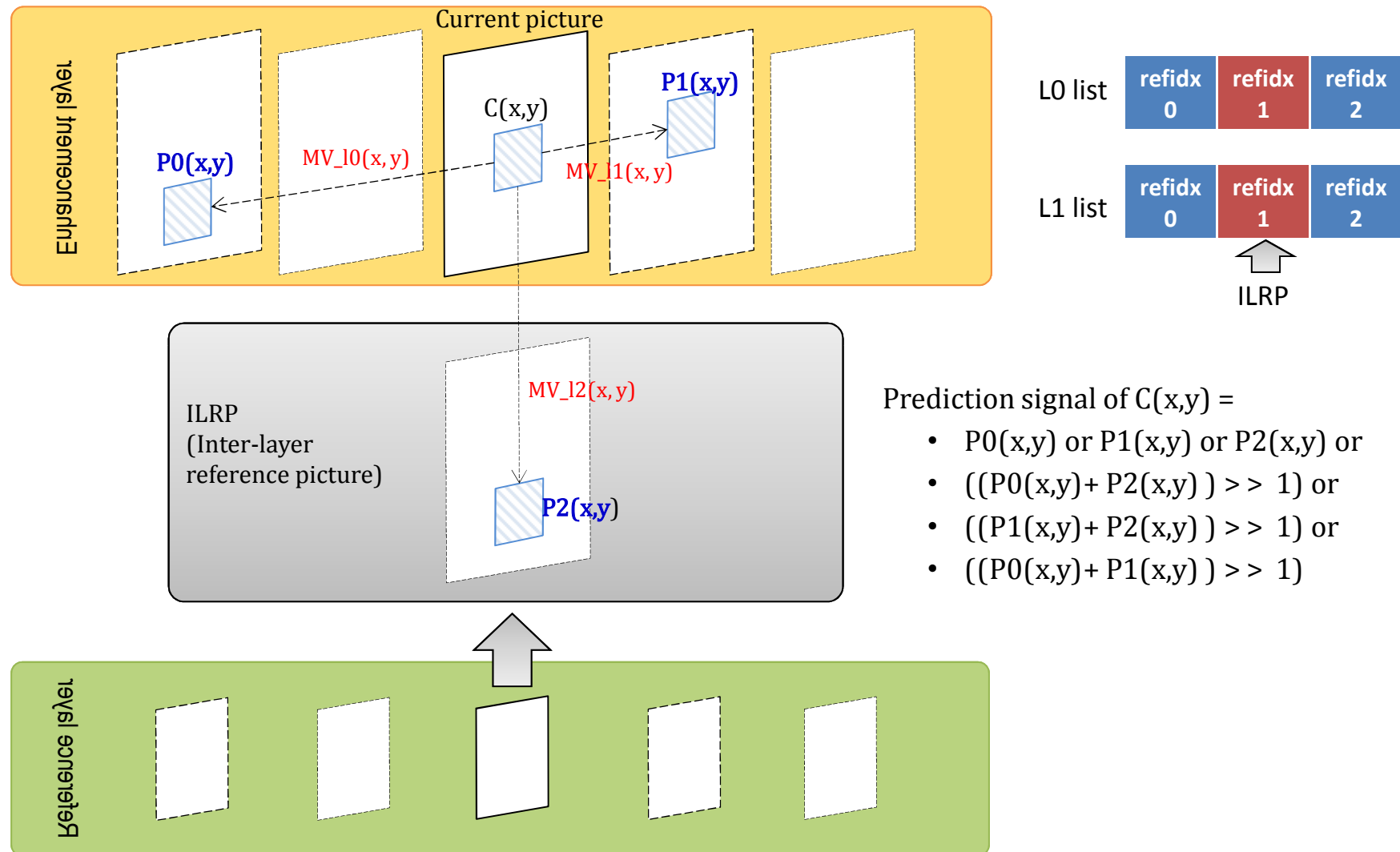
❑ The process for the reference picture list construction of the enhancement layer

1. An ILRP is generated
2. The reference picture lists, L0 and L1, are constructed as signaled in slice header of the enhancement layer
3. The ILRP is added into the reference picture lists of the enhancement layer



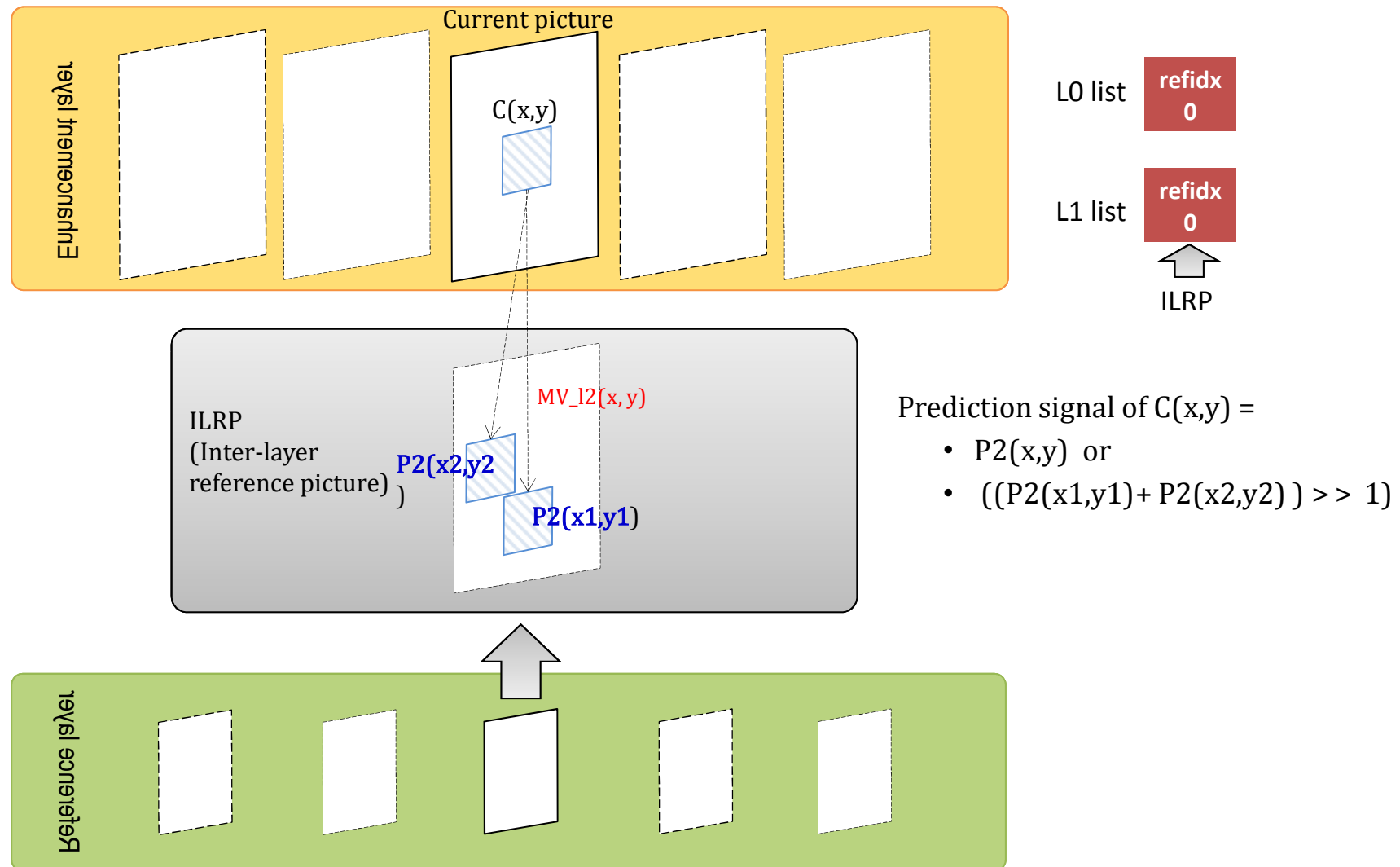
Inter-layer texture prediction

□ Prediction structure for non-random access picture



Inter-layer texture prediction

□ Prediction structure for random access picture



Inter-layer texture prediction

❑ Prediction modes used in the proposal

- ❖ Intra prediction mode: defined as in HEVC spec.
- ❖ Inter prediction mode: defined as in HEVC spec.
- ❖ Inter-layer prediction mode
 - Prediction derived only from the ILRP
- ❖ Combined inter-layer prediction mode
 - Prediction derived from ILRP and temporal reference picture of the enhancement layer

❑ Syntax for inter-layer texture prediction

- ❖ Syntax for B slice is used without any modification
 - Pictures in the enhancement layer are encoded as B frames except having ILRP in the reference picture lists

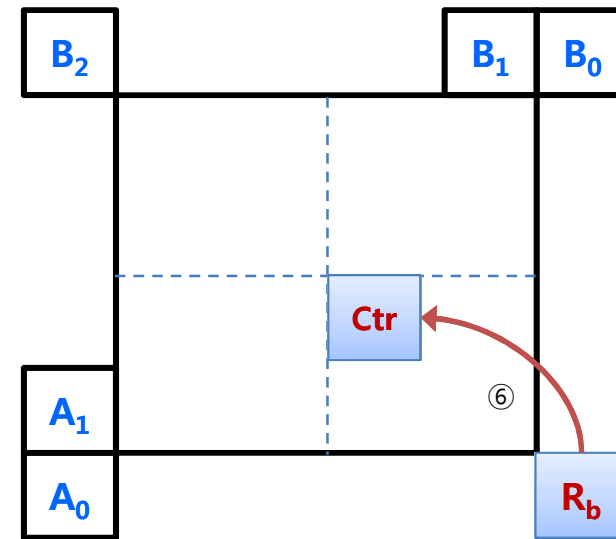
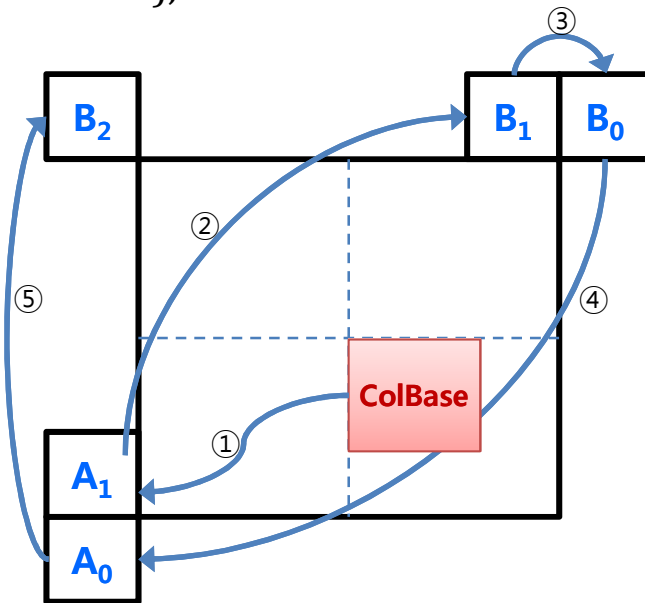
Inter-layer motion vector prediction

- ❑ **Motion information of the co-located PU in the reference layer**
 - ❖ Is added to merge/skip and motion vector prediction candidates
 - ❖ Can be applied to the single loop decoding design as well as multiple loop decoding design

Inter-layer motion vector prediction

❑ Merge/skip candidates

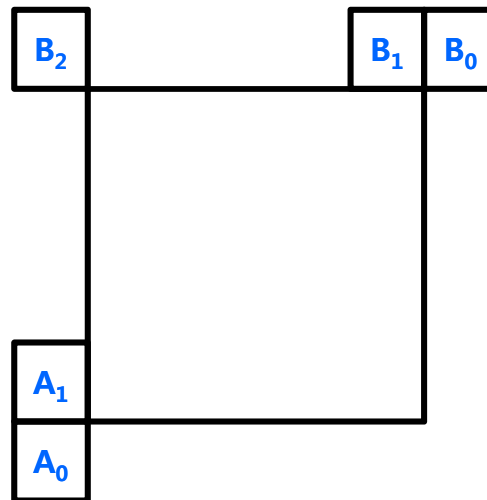
1. The motion information of the co-located PU in the reference layer (ColBase) is added as the first candidate out of 5 candidates.
2. Spatial merge candidates (A_1 , B_1 , B_0 , A_0 , B_2)
3. Temporal merge candidate (R_b , Ctr)
 - When the reference picture of temporal merge candidate or the co-located PU in temporally co-located picture is the corresponding picture of the reference layer (i.e. ILRP), it is set to be “unavailable”



Inter-layer motion vector prediction

❑ Motion vector prediction candidates

- ❖ When the current PU is predicted from the corresponding picture of the reference layer (i.e. ILRP)
 1. Spatial motion vector candidates (A_0, A_1, B_0, B_1, B_2)
 - A neighboring PU predicted from the ILRP is set to be “available”
 2. Temporal motion vector candidate is set to be “unavailable”

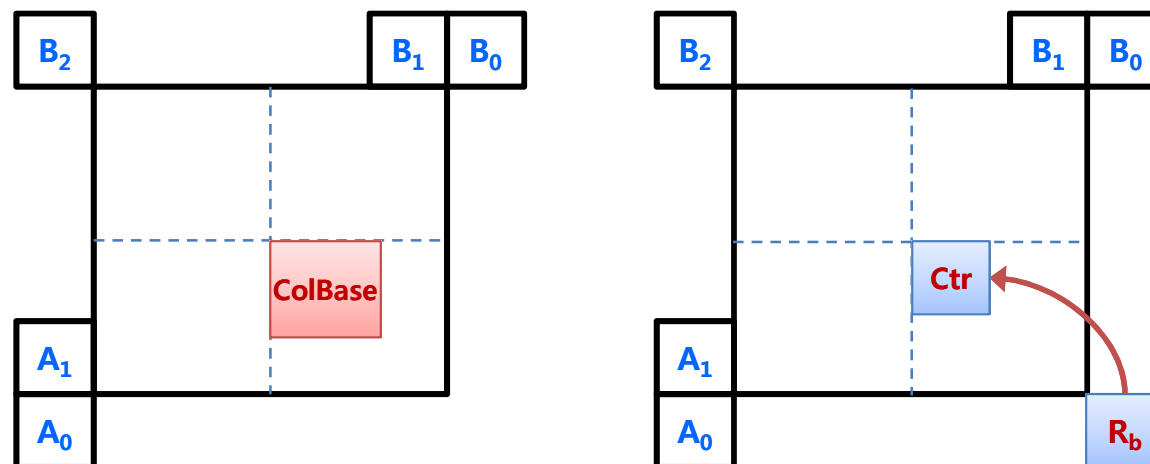


Inter-layer motion vector prediction

❑ Motion vector prediction candidates

❖ When the current PU is predicted from the temporal reference picture of the enhancement layer

1. The motion information of the co-located PU in the reference layer (ColBase) is added as the first candidate (after scaling if necessary).
2. Spatial motion vector candidates (A_0, A_1, B_0, B_1, B_2)
 - A neighboring PU predicted from the temporal reference picture of the enhancement layer(ILRP) is set to be “available”.
3. Temporal motion vector candidate (R_b , Ctr)
 - A temporal motion vector candidate whose reference picture is the corresponding picture of the reference layer(ILRP) is set to be “unavailable”.



Up-sampling filters

- ❑ For up-sampling the reconstructed picture of the reference layer
 - ❖ DCT-IF filters with 6 bits filter coefficients representation

Position	Luma filter coefficients	Chroma filter coefficients
1/3	{-1, 5, -12, 53, 26, -9, 4, -1}	{-5, 25, 51, -7}
1/2	{-1, 4, -11, 40, 40, -11, 4, -1}	{-4, 36, 36, -4}
2/3	{-1, 4, -9, 26, 53, -12, 5, -1}	{-7, 51, 25, -5}

Compression performance

□ Performance of EL-only

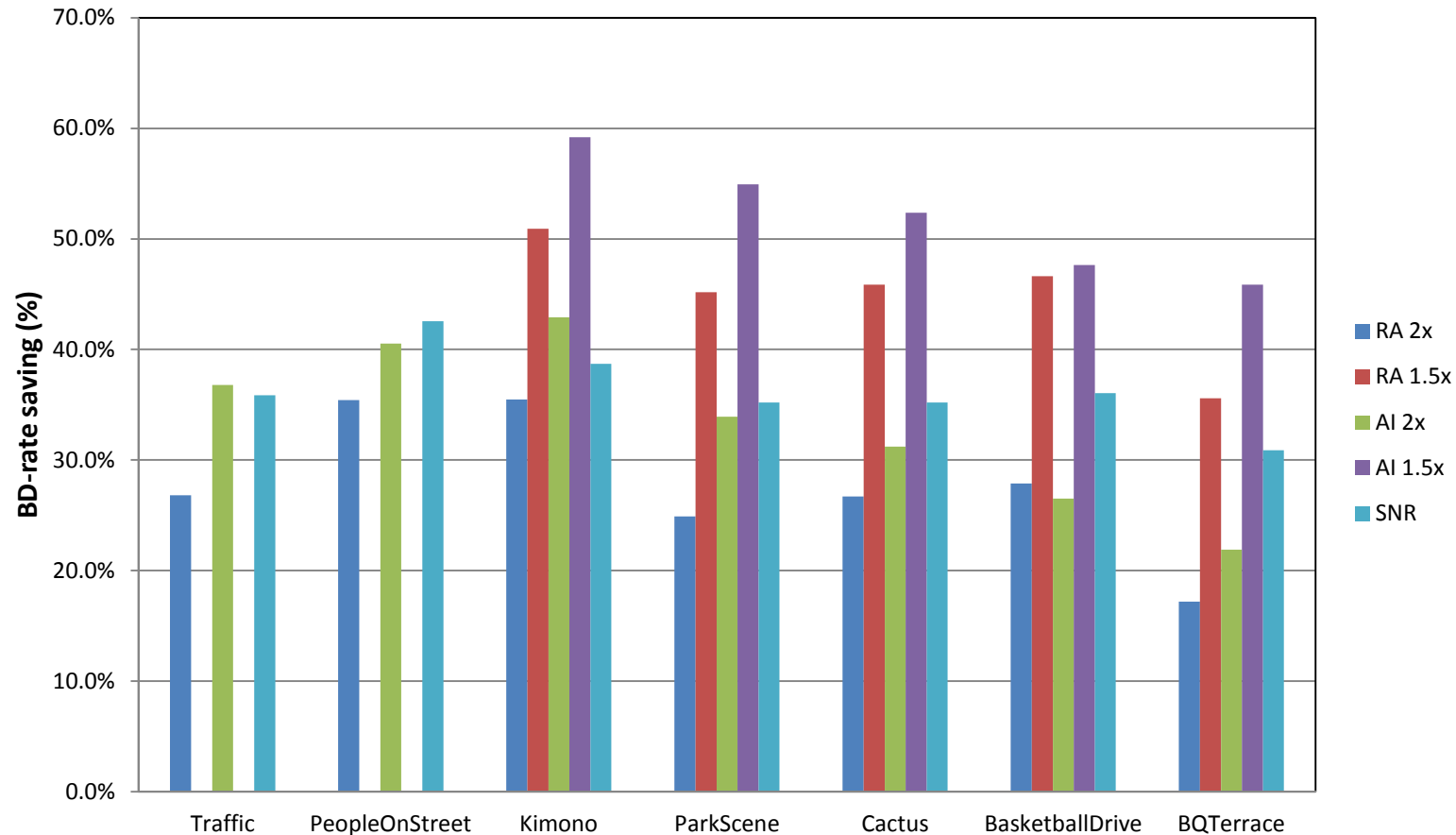
Class	Actual bitrate					Target bitrate				
	RA 2x	RA 1.5x	AI 2x	AI 1.5x	RA-SNR	RA 2x	RA 1.5x	AI 2x	AI 1.5x	RA-SNR
A ⁺	-31.1		-38.6		-39.2	-30.8		-38.6		-38.8
B	-26.4	-44.8	-31.3	-52.0	-35.2	-26.2	-44.7	-31.2	-51.9	-34.9
All	-28.8	-44.8	-35.0	-52.0	-37.2	-28.5	-44.7	-34.9	-51.9	-36.8

□ Performance of EL+BL

Class	Actual bitrate					Target bitrate				
	RA 2x	RA 1.5x	AI 2x	AI 1.5x	RA-SNR	RA 2x	RA 1.5x	AI 2x	AI 1.5x	RA-SNR
A ⁺	-20.8		-26.8		-25.9	-20.6		-26.8		-25.5
B	-17.7	-26.0	-21.5	-30.3	-23.2	-17.6	-25.9	-21.4	-30.3	-23.0
All	-19.3	-26.0	-24.2	-30.3	-24.5	-19.1	-25.9	-24.1	-30.3	-24.3

Compression performance

□ Summary



Complexity Analysis

❑ Encoding time analysis

❖ Performed with a cluster using 64-bit Linux

➤ 11 machines: Intel i7 990x Extreme, 6 Core, 3.46 GHz, 24GB RAM

❖ Encoding time relative to anchors (QP=22)

	RA 2x	RA 1.5x	AI 2x	AI 1.5x	RA-SNR
Encoding Time	1.28	1.43	2.62	2.68	1.93

❑ Decoding time analysis

❖ Performed with a PC using 64 bit Window 7

➤ Intel i7 990x Extreme, 6 Core, 3.46 GHz, 24GB RAM

❖ Decoding time relative to anchors (QP=22)

	RA 2x	RA 1.5x	AI 2x	AI 1.5x	RA-SNR
Decoding Time	1.41	1.67	1.31	1.49	1.94

Summary

☐ Minimum modifications on HEVC version 1 draft standard

- ❖ minor syntax changes
 - reserved_one_5bits to layer_id_plus1
 - inter_layer_prediction_flag in SPS for enabling inter-layer prediction.
- ❖ Inter-layer texture prediction
 - Using a reconstructed picture of the reference layer as the reference picture for the enhancement layer
- ❖ Modification on merge/skip & motion vector prediction candidates
 - to consider the corresponding PU on the reference layer
- ❖ ALF is enabled to the enhancement layer
- ❖ Up-sampling filters are developed

☐ Be easily extended to support multi-view scalability at the same time

☐ Can support coding standard scalability (AVC base layer)

☐ Performance

- ❖ EL only: 28.8% (RA 2x) ~ 52.0% (AI 1.5x) bitrate savings
- ❖ EL + BL: 19.3% (RA 2x) ~ 30.3%(AI 1.5x) bitrate savings
- ❖ Decoding time: 1.31 ~ 1.94 times relative to anchors