|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11**  11th Meeting: Shanghai, CN, 10-19 Oct., 2012 | Document: JCTVC-K0037 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Description of scalable video coding technology proposal by ETRI and Kwangwoon Univ.** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Jung Won Kang  Hahyun Lee Jinho Lee Jin Soo Choi Jin Woong Kim  Junghak Nam  Hyomin Choi Donggyu Sim | Tel: Email:  Tel: Email: | +82-42-860-5137 [jungwon@etri.re.kr](mailto:jungwon@etri.re.kr) [hanilee@etri.re.kr](mailto:hanilee@etri.re.kr) [jinosoul@etri.re.kr](mailto:jinosoul@etri.re.kr) [jschoi@etri.re.kr](mailto:jschoi@etri.re.kr) [jwkim@etri.re.kr](mailto:jwkim@etri.re.kr)  +82-2-940-5470 [qejixfyza@kw.ac.kr](mailto:qejixfyza@kw.ac.kr) [hyomin06@kw.ac.kr](mailto:hyomin06@kw.ac.kr) [dgsim@kw.ac.kr](mailto:dgsim@kw.ac.kr) |
|  |  |  |  |
| *Source:* | ETRI (Electronics and Telecommunications Research Institute) KWU (Kwangwoon University) | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

This document describes the scalable video coding technology proposal by ETRI and Kwangwoon Univ. The proposal is based on multiple loop decoding process which fully reconstructs the picture of all the layers with motion compensation for each layer. The proposal is developed to support mainly spatial scalability and SNR scalability. However, it is reported that the proposal can be extended to support multi-view scalability and coding standard scalability because of multiple loop decoding structure.

The proposal contains tools for enhancement layer. For inter-layer texture prediction, the reconstructed picture of the reference layer is added into the reference picture lists L0 and L1 for the corresponding picture of the enhancement layer. 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. DCT-IF filters are developed for up-sampling the reconstructed picture of the reference layer. In addition, ALF is applied to the enhancement layer.

For spatial scalability, compared to HEVC enhanced resolution single layer anchor, average BD-rate improvements are 28.8%(Y), 16.6%(U), and 15.2%(V) in case of spatial resolution of 2, and average BD-rate improvements are 44.8%(Y), 35.5%(U), and 33.3%(V) in case of spatial resolution of 1.5. For intra-only spatial scalability, compared to HEVC enhanced resolution single layer anchor, average BD-rate improvements are 35.0%(Y), 32.8%(U), and 32.5%(V) in case of spatial resolution of 2, and average BD-rate improvements are 52.0%(Y), 51.5%(U), and 51.6%(V) in case of spatial resolution of 1.5. For SNR scalability, compared to HEVC enhanced SNR single layer anchor, average BD-rate improvements are 37.2%(Y), 25.6%(U), and 22.2%(V).

# Contents

[Abstract 1](#_Toc336848983)

[Contents 2](#_Toc336848984)

[1 Introduction 3](#_Toc336848985)

[2 Algorithm description 3](#_Toc336848986)

[2.1 CTU/TU/PU partitioning 4](#_Toc336848987)

[2.2 Up-sampling filters 5](#_Toc336848988)

[2.3 Inter-layer prediction 7](#_Toc336848989)

[2.3.1 Inter-layer texture prediction 7](#_Toc336848990)

[2.3.2 Inter-layer motion vector prediction 9](#_Toc336848991)

[2.4 Inter-frame prediction 10](#_Toc336848992)

[2.5 Intra-frame prediction 10](#_Toc336848993)

[2.6 In-loop filtering 11](#_Toc336848994)

[2.6.1 De-blocking filter 11](#_Toc336848995)

[2.6.2 SAO 11](#_Toc336848996)

[2.6.3 ALF 11](#_Toc336848997)

[2.7 Spatial transforms 11](#_Toc336848998)

[2.8 Quantization 11](#_Toc336848999)

[2.9 Motion/mode/parameter coding 11](#_Toc336849000)

[2.9.1 Merge/Skip 11](#_Toc336849001)

[2.9.2 Motion vector coding 11](#_Toc336849002)

[2.9.3 Intra prediction mode coding 11](#_Toc336849003)

[2.9.4 SAO parameter coding 11](#_Toc336849004)

[2.10 Entropy coding of residual / transform coefficients 11](#_Toc336849005)

[2.11 High-layer syntax 11](#_Toc336849006)

[3 Syntax and semantics description 11](#_Toc336849007)

[4 Compression performance discussion 12](#_Toc336849008)

[4.1 Category 1 (HEVC base layer) 12](#_Toc336849009)

[4.1.1 Spatial scalability 12](#_Toc336849010)

[4.1.2 Intra-only spatial scalability 13](#_Toc336849011)

[4.1.3 SNR scalability 15](#_Toc336849012)

[4.2 Category 2 (AVC base layer) 15](#_Toc336849013)

[4.2.1 Spatial scalability 15](#_Toc336849014)

[4.2.2 Overall 15](#_Toc336849015)

[5 Complexity analysis 16](#_Toc336849016)

[5.1 Encoding time and measurement methodology 16](#_Toc336849017)

[5.2 Decoding time and measurement methodology and comparison vs. anchor bitstreams decoded by HM 17](#_Toc336849018)

[5.3 Description of computing platform used to determine encoding and decoding times reported in sections 5.1 and 5.2 18](#_Toc336849019)

[5.4 Expected memory usage of encoder 18](#_Toc336849020)

[5.5 Expected memory usage of decoder 18](#_Toc336849021)

[5.6 Complexity characteristics of encoder motion estimation and partitioning selection in enhancement layer(s) 18](#_Toc336849022)

[5.7 Complexity characteristics of decoder motion compensation in enhancement layer(s) 19](#_Toc336849023)

[5.8 Complexity characteristics of encoder intra-frame prediction type and partitioning selection in enhancement layer(s) 19](#_Toc336849024)

[5.9 Complexity characteristics of decoder intra-frame prediction operation in enhancement layer(s) 19](#_Toc336849025)

[5.10 Complexity characteristics of upsampling filters and transforms specific in enhancement layer(s) 19](#_Toc336849026)

[5.11 Complexity characteristics of quantization and inverse quantization in enhancement layer(s) 19](#_Toc336849027)

[5.12 Complexity characteristics of encoder entropy coding operation in enhancement layer(s) 19](#_Toc336849028)

[5.13 Complexity characteristics of decoder entropy decoding operation in enhancement layer(s) 19](#_Toc336849029)

[5.14 Degree of capability for encoder parallel processing 19](#_Toc336849030)

[5.15 Degree of capability for decoder parallel processing 20](#_Toc336849031)

[6 Software implementation description 20](#_Toc336849032)

[7 References 20](#_Toc336849033)

[8 Patent rights declaration(s) 20](#_Toc336849034)

# Introduction

This document describes the scalable video coding technology proposal by ETRI and Kwangwoon Univ. The proposal is based on multiple loop decoding process which fully reconstructs the picture of all the layers with motion compensation for each layer.

Basically, the single-loop design does not reconstruct a reference layer for the inter-slice cases. It can decrease decoder complexity by skipping motion compensation. On the other hands, the multi-loop decoding design fully reconstructs the frames of all the layers with motion compensation for each layer. The multiple loop decoding design is known to be efficient in coding gain while requiring more computational complexity than the single-loop scalable structure. In addition, the multiple loop design could allow us to easily support multi-view scalability at the same time, since multiple loop scalable structure which perfectly reconstruct frame of all layer can display any views of the multi-view configuration as view scalability.

The proposal is developed to support mainly spatial scalability and SNR scalability, but can support coding standard scalability because of multiple loop decoding structure.

The following describes the 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

🞏 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)

# Algorithm description

Encoder and Decoder block diagram of the proposal are described in Figure 1 and Figure 2. Layer 0 which is a base layer can be decoded with HEVC HM6.1, with the requirement of the backward compatibility as described in CfP document [1]. The layer 1 which is an enhancement layer is coded by HEVC tools and additional inter-layer prediction tools represented by red blocks in Figure 1 and Figure 2. Features applied to the proposal are as follows.

* **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. For the proposed inter-layer texture prediction, the reconstructed picture of the reference layer, whose spatial resolution is the same as that of the enhancement layer, is added into the reference picture lists, L0 and L1, for the corresponding picture of the enhancement layer although it is not a temporal reference 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 filter**: Interpolation filter defined in HEVC draft spec[2] is used for up-sampling filter of luma and chroma samples in case of spatial resolution factor(enhancement/base) of 2 and the 8 and 4 taps DCI-IF filters are designed for up-sampling of luma and chroma samples, respectively, in case of spatial resolution factor of 1.5.

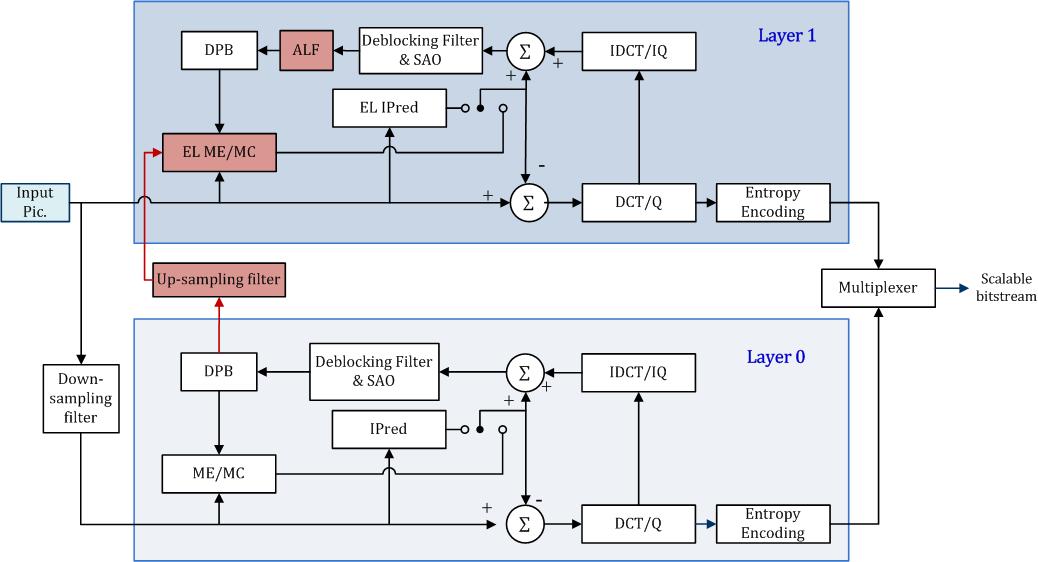


Figure 1. Encoder Block Diagram



Figure 2. Decoder Block Diagram

## CTU/TU/PU partitioning

CTU/TU/PU partitioning is used as specified in HEVC draft spec[2].

## Up-sampling filters

Since the proposed codec is based on the multiple loop decoding design, the reconstructed picture of the reference layer is up-sampled when the spatial resolution of the reference layer is different to that of the enhancement layer. The proposal employs the DCT-IF filter for up-sampling the reconstructed picture of the reference layer. Table 1 presents up-sampling filter coefficients for luma samples, which is the 8-taps DCT-IF filter with 6 bits filter coefficients representation.

Table 1. Up-sampling filter coefficients for luma samples

|  |  |
| --- | --- |
| Position | Filter coefficients |
| 1/2 | {-1, 4, -11, 40, 40, -11, 4, -1} |
| 1/3 | {-1, 5, -12, 53, 26, -9, 4, -1} |
| 2/3 | {-1, 4, -9, 26, 53, -12, 5, -1} |

In case of spatial resolution factor (enhancement/base) of 2, Interpolation filter defined in HEVC draft spec[2] is used for the up-sampling filter for luma samples. Figure 3 shows an example of up-sampling in case of the spatial resolution factor of 2. Pixel values of the enhancement layer are derived as follows.

* Pixels filled with the same color indicated by “A”: duplicate the pixel values of the base layer.
* Pixels filled with the same color indicated by “B”: calculated by applying 1/2 position filter coefficients to vertical or horizontal direction.
* Pixels filled with the same color indicated by “C”: calculated by applying 1/2 position filter coefficients to vertical and horizontal directions.

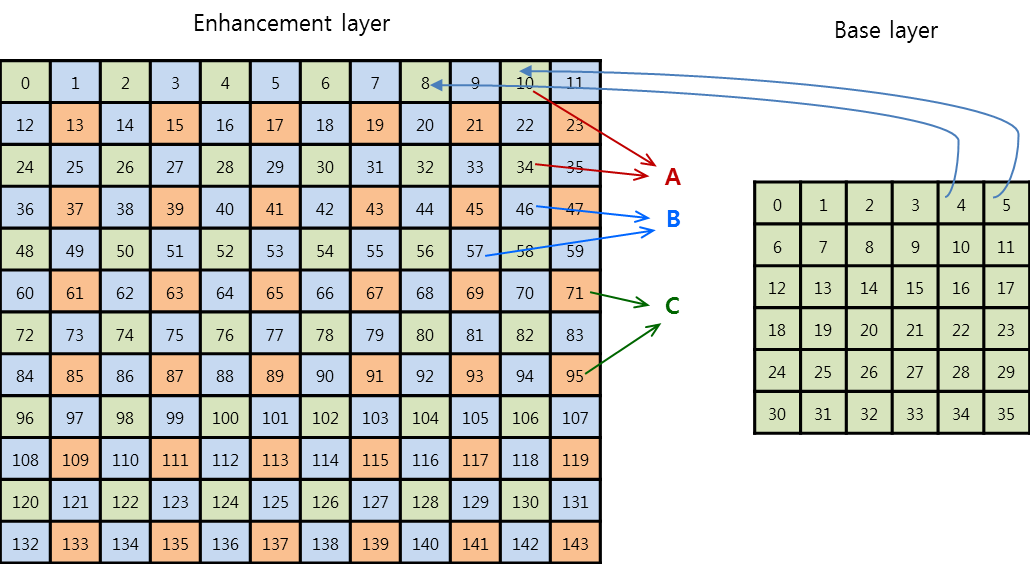


Figure 3. An example of up-sampling in case of spatial resolution factor of 2

In case of the spatial resolution factor of 1.5, 8-taps DCI-IF filters developed with 6 bits filter coefficients representation for up-sampling of luma samples. Figure 4 shows an example of up-sampling in case of spatial resolution factor of 1.5. Pixel values of the enhancement layer are derived as follows.

* Pixels filled with the same color indicated by “A”: duplicate the pixel values of the base layer, whose positions are in odd rows and odd columns.
* Pixels filled with the same color indicated by “B”: calculated by applying 1/3 position filter coefficients to vertical or horizontal direction.
* Pixels filled with the same color indicated by “C”: calculated by applying 2/3 position filter coefficients to vertical or horizontal direction.
* Pixels filled with the same color indicated by “D”: calculated by applying 1/3 position filter coefficients to vertical and horizontal directions.
* Pixels filled with the same color indicated by “E”: calculated by applying 2/3 position filter coefficients to vertical direction and 1/3 position filter coefficients to horizontal direction, or calculated by applying 1/3 position filter coefficients to vertical direction and 2/3 position filter coefficients to horizontal direction.
* Pixels filled with the same color indicated by “F”: calculated by applying 2/3 position filter coefficients to vertical and horizontal directions.

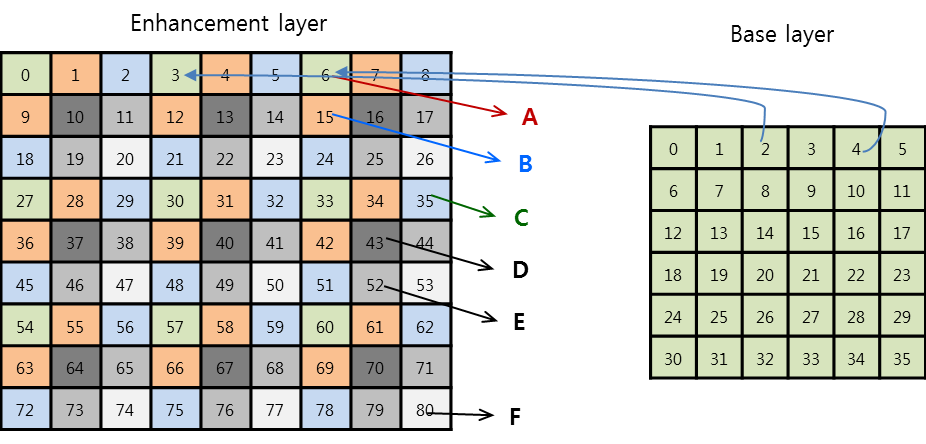


Figure 4. An example of up-sampling in case of spatial resolution factor of 1.5

Table 2 describes up-sampling filter coefficients for chroma samples, which is the 4-taps DCT-IF filter with 6 bits filter coefficients representation. The filtering process for chroma samples is the same as the filtering process for luma samples.

Table 2. Up-sampling filter coefficients for chroma samples

|  |  |
| --- | --- |
| Position | Filter coefficients |
| 1/2 | {-4, 36, 36, -4} |
| 1/3 | {-5, 25, 51, -7} |
| 2/3 | {-7, 51, 25, -5} |

## Inter-layer prediction

### Inter-layer texture prediction

In this proposal, a picture of the reference layer is completely reconstructed, up-sampled (if necessary), and inserted into the reference picture lists for the corresponding picture of the enhancement layer. This reconstructed picture of the reference layer, whose spatial resolution is the same as that of pictures in the enhancement layer, is defined as inter-layer reference picture (ILRP). Since ILRP is added into the reference picture lists, L0 and L1, the number of the reference pictures in L0 and L1 increases by 1, respectively, compared to the main configuration of HEVC HM 6.1 [3].

In detail, the process for reference picture list construction of the enhancement layer is as follows.

1. An ILRP is generated by reconstructing and up-sampling (if necessary) the corresponding picture of the reference layer.
2. The reference picture lists, L0 and L1, are constructed as signaled in slice header of the enhancement layer.
3. The ILRP, that is assigned to the reference index of “1”, is added into the reference picture lists, L0 and L1 of the enhancement layer while reference index of reference picture whose reference index was originally greater than “1” is increased by 1 if there is any.
4. Reference picture list modification can be performed if necessary. (In the proposal, it is not utilized.)

After the reference picture lists are constructed, motion estimation and compensation process that are the same as described in the HEVC draft spec[2] shall be performed. It should be noted that, since the ILRP is used as a reference picture of the enhancement layer and included in both reference picture lists, L0 and L1, at the enhancement layer, I and P frames can be removed from enhancement layer and all the pictures in the enhancement layer are encoded as B frame.

In addition, by allowing the reconstructed picture of the reference layer (i.e. ILRP) to be used as a reference picture for the corresponding picture of the enhancement layer, the enhancement layer picture can have a prediction structure to be predicted from not only from co-located region of the reference layer but also other region of the reference layer with a motion vector, and to be combined(bi-) predicted from temporal reference picture of the enhancement layer and ILRP. The latter is defined as the combined inter-layer prediction mode.

Therefore, conceptually, prediction mode used in the proposal can be categorized as follows.

* **Intra prediction mode**: same as described in HEVC draft spec[2].
* **Inter prediction mode**: same as described in HEVC draft spec[2]. There can be uni or bi-inter prediction.
* **Inter-layer prediction mode**: prediction derived only from the reconstructed picture of the reference layer(i.e. ILRP). There can be uni or bi-inter-layer prediction.
* **Combined inter-layer prediction mode**: prediction derived from temporal reference picture of the enhancement layer and the reconstructed picture of the reference layer(i.e. ILRP).

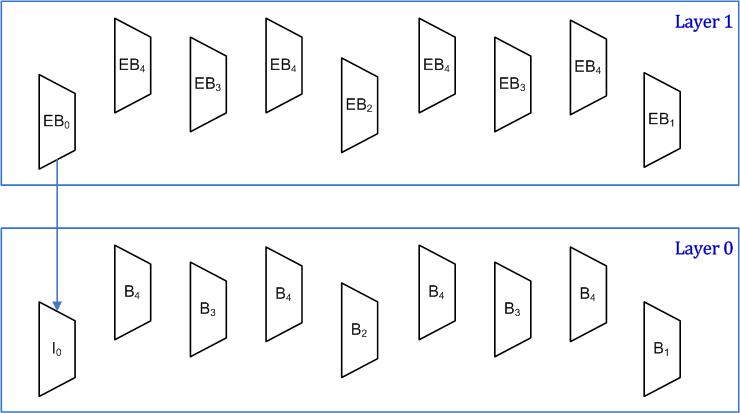
Despite of the detailed explanation above, it should be noted that, except inserting the ILRP into the reference picture lists L0 and L1 of the enhancement layer, slice of the enhancement layer utilizes exactly the same prediction scheme and syntax as B slice of the base layer for inter-prediction.

In Figure 5, examples of the prediction structure of the proposal are described.

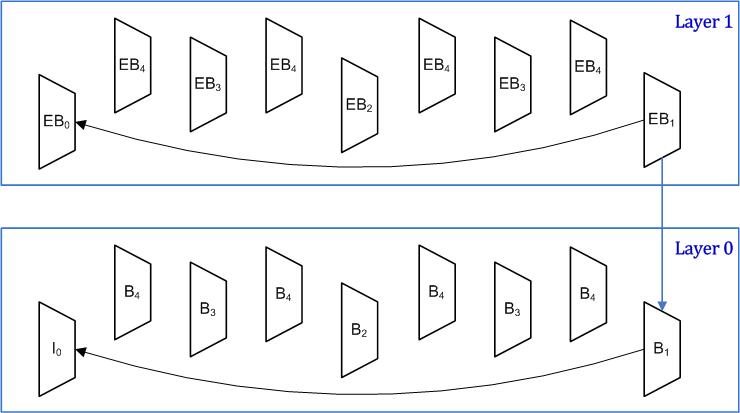
A random access picture, EB0, whose reference layer is coded as I slice can be coded in intra prediction mode, or uni- or bi-inter-layer prediction mode. For the inter-layer prediction mode, the ILRP is assigned to the reference index of “0” in both reference picture lists L0 and L1 at the enhancement layer because there is no other reference picture in EB0. Therefore, EB0 picture can be uni- or bi-predicted from the ILRP as described in Figure 5(a).

EB1 picture whose reference layer is coded as B slice (i.e. GPB slice) can be coded in intra prediction mode, uni- or bi-inter prediction mode, uni- or bi-inter-layer prediction mode, or combined inter-layer prediction mode. For the inter-layer prediction mode, the ILRP is assigned to the reference index of “1” in both the reference picture lists L0 and L1 at the enhancement layer. Therefore, EB1 picture can be uni- or bi-predicted from only EB0 picture or only the ILRP, or bi-predicted from EB0 picture and the ILRP as described in Figure 5(b).

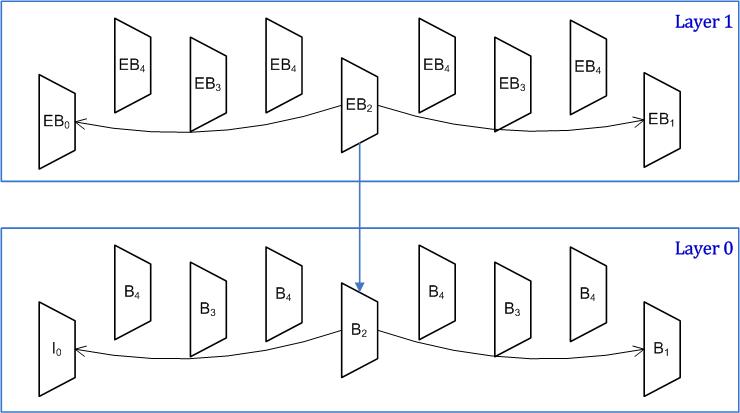
EB2 picture whose reference layer is coded as B slice can be coded in intra prediction mode, uni- or bi-inter prediction mode, uni- or bi-inter-layer prediction mode, or combined inter-layer prediction mode. For the inter-layer prediction mode, the ILRP is assigned to the reference index of “1” in both the reference picture lists L0 and L1 at the enhancement layer. Therefore, EB2 picture can be uni- or bi-predicted from only EB0 picture, only EB1 picture, or only the ILRP, or bi-predicted from any combination of EB0, EB1 and the ILRP as described in Figure 5(c). The EB3 and EB4 pictures are coded in the same way as the EB2 picture.



(a) Prediction structure for an enhancement layer picture that can be compressed in intra prediction mode or inter-layer prediction mode



(b) Prediction structure for an enhancement layer picture that can be compressed in intra prediction mode, inter prediction mode, inter-layer prediction mode, or combined inter-layer prediction mode



(c) Prediction structure for an enhancement layer picture that can be compressed in intra prediction mode, inter prediction mode, inter-layer prediction mode, or combined inter-layer prediction mode

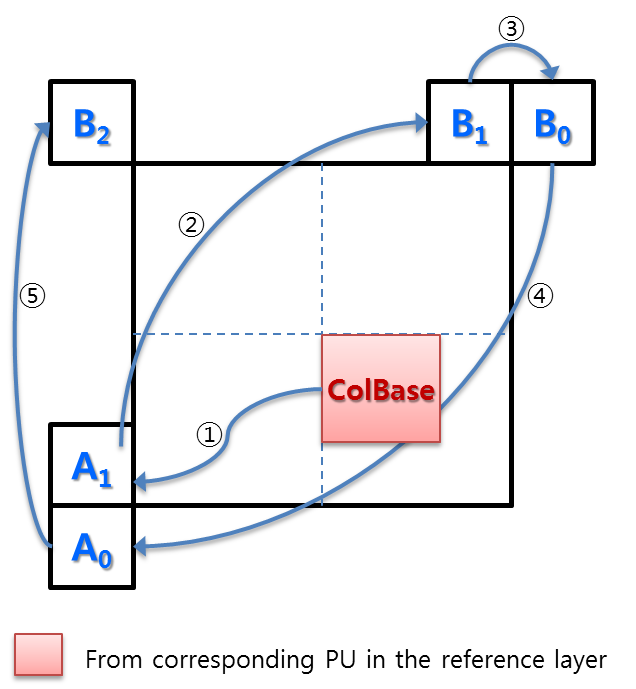
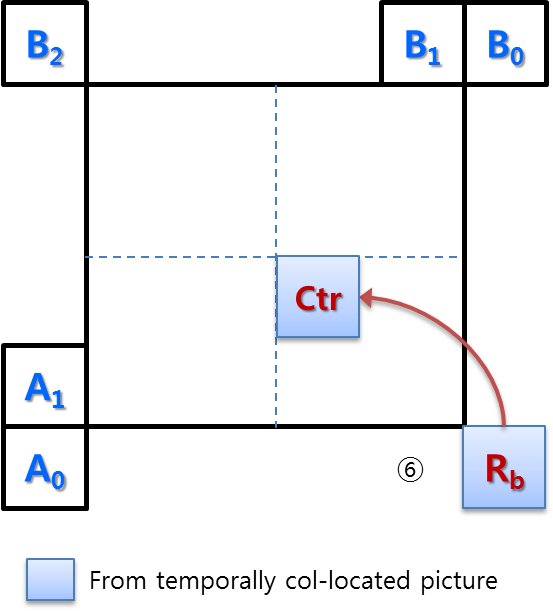
Figure 5. Examples of the prediction structure

### Inter-layer motion vector prediction

Since motion information of the current PU in the enhancement layer is likely to be similar to the co-located PU in the reference layer, motion information of the co-located PU in the reference layer is employed as one of candidates for merge/skip mode and motion vector prediction. When a motion vector of the co-located PU in the reference layer is used as merge/skip or motion vector prediction candidate, it is scaled by the spatial resolution factor.

It should be noted that this method can be applied to single loop decoding design as well as multiple loop decoding design.

#### Merge/skip candidates

(a) (b)

Figure 6. Positions of merge candidates

In the proposal, there is modification in constructing of merge candidates as described in Figure 6.

Since there is high correlation between motion information of the co-located PU in the reference layer and the current PU in the enhancement layer, the motion information of the co-located PU in the reference layer is added as the first merge candidate.

The numbers in Figure 6 show the order of derivation of merge candidates. In the derivation of merge candidates, maximum four merge candidates are selected among that are located in the position as described in Figure 6(a) and, then, the temporal merge candidate is selected between two candidate positions as described in Figure 6(b). In derivation of temporal merge candidate, 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), the temporal merge candidate is considered to be unavailable. The remaining part of construction process for merge candidates is the same as described in HEVC draft spec[2].

#### Motion vector prediction candidates

Basically, derivation process for motion vector prediction candidate is the same as described in HEVC draft spec[2] except modification described as follows.

Since the ILRP is inserted into reference picture lists at the enhancement layer, a PU can be predicted from the corresponding picture of the reference layer, i.e. ILRP, or from the previously encoded picture of the enhancement layer, i.e. temporal reference picture. According to the type of current PU’s reference picture, the derivation process for motion vector prediction candidate list should be different.

When the current PU is predicted from the corresponding picture of the reference layer, PU, among PUs located in five different positions described in Figure 7, which is predicted from the corresponding picture of the reference layer, is only considered as available for motion vector prediction candidates. In addition, temporal motion vector candidate is set to be unavailable.

When the current PU is predicted from the temporal reference picture of the enhancement layer, the motion information of the co-located PU in the reference layer is added as the first motion vector prediction candidate. If POC of the current PU’s reference picture is different from that of the co-located PU’s reference picture in the reference layer, motion vector of the co-located PU is scaled as the same manner of temporal scaling as described in HEVC draft spec[2]. In addition, among PUs located in five different positions described in Figure 7, PU which is predicted from the temporal reference picture of the enhancement layer is only considered to be available for motion vector prediction candidates. Temporal motion vector candidate whose reference picture is the corresponding picture of the reference layer is also set to be unavailable. The remaining part of construction process for merge candidates is the same as described in HEVC draft spec[2].

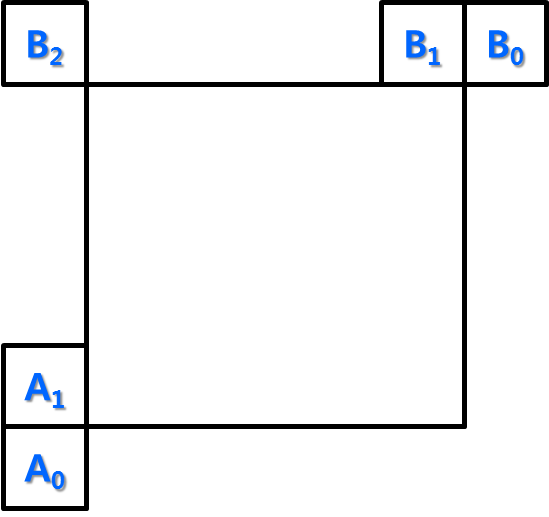


Figure 7. Positions of spatial motion vector prediction candidates

## Inter-frame prediction

Inter-frame prediction is used as specified in HEVC draft spec[2].

## Intra-frame prediction

Inter-frame prediction is used as specified in HEVC draft spec[2].

## In-loop filtering

### De-blocking filter

De-blocking filter is used as specified in HEVC draft spec[2].

### SAO

SAO is used as specified in HEVC draft spec[2].

### ALF

ALF is used as specified in HEVC draft spec[2] for enhancement layer.

## Spatial transforms

Spatial transforms are used as specified in HEVC draft spec[2].

## Quantization

Quantization process is used as specified in HEVC draft spec[2].

## Motion/mode/parameter coding

### Merge/Skip

Merge/skip candidates are modified as described in 2.3.2.1 merge/skip candidates.

### Motion vector coding

Motion vector prediction candidates are modified as described in 2.3.2.2 motion vector prediction candidates.

### Intra prediction mode coding

Intra prediction mode coding is used as specified in HEVC draft spec[2]

### SAO parameter coding

SAO parameter coding is used as specified in HEVC draft spec[2]

## Entropy coding of residual / transform coefficients

CABAC (Context-Adaptive Binary Arithmetic Coding) is used as specified in HEVC draft spec[2].

## High-layer syntax

High-level syntax as specified in HEVC draft spec[2] is used except “reserved\_one\_5bits” in the NAL unit header is replaced by “layer\_id\_plus1” indicating NAL unit for enhancement layer. No additional NAL unit type or slice type is defined. A NAL unit, when “layer\_id\_plus1” specified in NAL unit header is greater than 1, is for enhancement layer such as SPS, PPS, and a coded slice for enhancement layer. In SPS for enhancement layer, sps\_extension\_flag is set to be “1” and a flag is defined in the extension field to indicate if inter-layer prediction is utilized in the bitstream.

# Syntax and semantics description

High-level syntax as specified in HEVC draft spec[2] is used except “reserved\_one\_5bits” in the NAL unit header is replaced by “layer\_id\_plus1” specifying the identifier for spatial scalability. In addition, no new syntax and semantics are introduced for the enhancement layer. Syntax and semantics for the enhancement layer are shared with the base layer specified in HEVC draft spec[2].

It should be noted that every coded slice of enhancement layer is assigned to slice type B since a coded slice of the enhancement layer can have at most two motion vectors for inter-layer prediction and combined inter-layer prediction as well as inter prediction.

In order to differentiate these three kinds of prediction modes i.e., inter-layer prediction mode, combined prediction mode, inter prediction mode, two more slice types for inter-layer prediction and combined inter-layer prediction could be defined later. However, those are not defined in this proposal since they need the same syntax elements for specifying the related motion information..

However, in order to indicate spatial layer switching point, a new NAL unit type may be necessary for the enhancement layer.

# Compression performance discussion

## Category 1 (HEVC base layer)

### Spatial scalability

The proposed enhanced resolution layers were created with a QP change (increasing or decreasing QP value by 1) in the middle of the sequence to match the target bitrate more accurately.

The followings are performance for the case of spatial resolution factor of 2.

Table 3. BD-rate for the proposed enhanced resolution layer relative to HEVC enhanced resolution single layer anchors (spatial resolution factor of 2)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **A+** | 3840x2048 | Traffic | -26.8 | -21.6 | -20.4 | -26.6 | -21.3 | -20.1 |
| 3840x2160 | PeopleOnStreet | -35.4 | -11.6 | -12.8 | -35.0 | -11.2 | -12.4 |
| **B** | 1920x1080 | Kimono | -35.5 | -26.1 | -23.7 | -35.3 | -25.9 | -23.5 |
| ParkScene | -24.9 | -18.8 | -19.3 | -24.6 | -18.5 | -19.0 |
| Cactus | -26.7 | -17.8 | -13.7 | -26.5 | -17.5 | -13.5 |
| BasketballDrive | -27.9 | -13.9 | -14.3 | -27.7 | -13.6 | -14.0 |
| BQTerrace | -17.2 | -6.7 | 2.1 | -17.0 | -6.4 | 2.4 |
| Summary | | | | | | | | |
| **A+** |  |  | -31.1 | -16.6 | -16.6 | -30.8 | -16.2 | -16.3 |
| **B** |  |  | -26.4 | -16.7 | -13.8 | -26.2 | -16.4 | -13.5 |
| **All** |  |  | -28.8 | -16.6 | -15.2 | -28.5 | -16.3 | -14.9 |

Table 4. BD-rate for the proposed enhanced resolution layer relative to HEVC enhanced resolution single layer anchor including base layer anchor (spatial resolution factor of 2)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **A+** | 3840x2048 | Traffic | -17.5 | -13.2 | -12.2 | -17.3 | -13.0 | -12.0 |
| 3840x2160 | PeopleOnStreet | -24.2 | -4.0 | -5.0 | -23.9 | -3.8 | -4.7 |
| **B** | 1920x1080 | Kimono | -24.6 | -16.4 | -14.4 | -24.5 | -16.2 | -14.3 |
| ParkScene | -16.4 | -11.2 | -11.8 | -16.3 | -11.0 | -11.6 |
| Cactus | -17.7 | -10.4 | -6.8 | -17.5 | -10.2 | -6.6 |
| BasketballDrive | -18.7 | -6.8 | -7.2 | -18.5 | -6.7 | -7.0 |
| BQTerrace | -11.2 | -1.7 | 5.9 | -11.1 | -1.5 | 6.2 |
| Summary | | | | | | | | |
| **A+** |  |  | -20.8 | -8.6 | -8.6 | -20.6 | -8.4 | -8.4 |
| **B** |  |  | -17.7 | -9.3 | -6.8 | -17.6 | -9.1 | -6.6 |
| **All** |  |  | -19.3 | -9.0 | -7.7 | -19.1 | -8.7 | -7.5 |

The followings are performance for the case of spatial resolution factor of 1.5.

Table 5. BD-rate for the proposed enhanced resolution layer relative to HEVC enhanced resolution single layer anchors (spatial resolution factor of 1.5)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **B** | 1920x1080 | Kimono | -50.9 | -44.0 | -41.6 | -50.7 | -43.7 | -41.4 |
| ParkScene | -45.2 | -38.9 | -39.7 | -45.0 | -38.7 | -39.5 |
| Cactus | -45.9 | -36.9 | -33.5 | -45.7 | -36.7 | -33.3 |
| BasketballDrive | -46.6 | -34.8 | -34.2 | -46.5 | -34.6 | -34.0 |
| BQTerrace | -35.6 | -22.9 | -17.4 | -35.4 | -22.7 | -17.2 |
| Summary | | | | | | | | |
| **B** |  |  | -44.8 | -35.5 | -33.3 | -44.7 | -35.3 | -33.1 |

Table 6. BD-rate for the proposed enhanced resolution layer relative to HEVC enhanced resolution single layer anchor including base layer anchor (spatial resolution factor of 1.5)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **B** | 1920x1080 | Kimono | -30.7 | -23.8 | -21.8 | -30.5 | -23.7 | -21.7 |
| ParkScene | -25.9 | -20.2 | -21.2 | -25.8 | -20.1 | -21.1 |
| Cactus | -26.3 | -18.7 | -15.7 | -26.2 | -18.6 | -15.6 |
| BasketballDrive | -27.3 | -16.8 | -16.2 | -27.2 | -16.7 | -16.1 |
| BQTerrace | -20.0 | -8.7 | -4.7 | -19.8 | -8.6 | -4.5 |
| Summary | | | | | | | | |
| **B** |  |  | -26.0 | -17.6 | -15.9 | -25.9 | -17.5 | -15.8 |

### Intra-only spatial scalability

The proposed enhanced resolution layers were created with a QP change (increasing or decreasing QP value by 1) in the middle of the sequence to match the target bitrate more accurately.

The followings are performance for the case of spatial resolution factor of 2.

Table 7. BD-rate for the proposed enhanced resolution layer relative to HEVC enhanced resolution single layer anchors (spatial resolution factor of 2)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **A+** | 3840x2048 | Traffic | -36.8 | -35.7 | -36.0 | -36.7 | -35.7 | -36.0 |
| 3840x2160 | PeopleOnStreet | -40.5 | -37.4 | -37.7 | -40.5 | -37.4 | -37.6 |
| **B** | 1920x1080 | Kimono | -42.9 | -40.6 | -41.4 | -42.9 | -40.6 | -41.3 |
| ParkScene | -33.9 | -32.8 | -33.3 | -33.9 | -32.7 | -33.2 |
| Cactus | -31.2 | -28.7 | -25.2 | -31.2 | -28.7 | -25.2 |
| BasketballDrive | -26.5 | -22.4 | -22.7 | -26.5 | -22.3 | -22.6 |
| BQTerrace | -21.9 | -20.6 | -18.5 | -21.9 | -20.5 | -18.5 |
| Summary | | | | | | | | |
| **A+** |  |  | -38.6 | -36.6 | -36.8 | -38.6 | -36.5 | -36.8 |
| **B** |  |  | -31.3 | -29.0 | -28.2 | -31.2 | -29.0 | -28.2 |
| **All** |  |  | -35.0 | -32.8 | -32.5 | -34.9 | -32.7 | -32.5 |

Table 8. BD-rate for the proposed enhanced resolution layer relative to HEVC enhanced resolution single layer anchor including base layer anchor (spatial resolution factor of 2)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **A+** | 3840x2048 | Traffic | -25.3 | -24.2 | -24.5 | -25.3 | -24.2 | -24.5 |
| 3840x2160 | PeopleOnStreet | -28.3 | -25.3 | -25.5 | -28.3 | -25.3 | -25.5 |
| **B** | 1920x1080 | Kimono | -30.6 | -28.3 | -29.1 | -30.6 | -28.2 | -29.1 |
| ParkScene | -23.9 | -22.9 | -23.5 | -23.9 | -22.9 | -23.5 |
| Cactus | -21.2 | -19.2 | -16.1 | -21.2 | -19.2 | -16.1 |
| BasketballDrive | -17.2 | -13.7 | -14.0 | -17.2 | -13.7 | -14.0 |
| BQTerrace | -14.3 | -13.2 | -11.6 | -14.3 | -13.2 | -11.5 |
| Summary | | | | | | | | |
| **A+** |  |  | -26.8 | -24.8 | -25.0 | -26.8 | -24.7 | -25.0 |
| **B** |  |  | -21.5 | -19.5 | -18.8 | -21.4 | -19.4 | -18.8 |
| **All** |  |  | -24.2 | -22.1 | -21.9 | -24.1 | -22.1 | -21.9 |

The followings are performance for the case of spatial resolution factor of 1.5.

Table 9. BD-rate for the proposed enhanced resolution layer relative to HEVC enhanced resolution single layer anchor (spatial resolution factor of 1.5)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **B** | 1920x1080 | Kimono | -59.2 | -56.1 | -57.4 | -59.2 | -56.1 | -57.3 |
| ParkScene | -54.9 | -54.5 | -55.6 | -54.9 | -54.5 | -55.5 |
| Cactus | -52.4 | -51.6 | -51.0 | -52.2 | -51.5 | -50.9 |
| BasketballDrive | -47.6 | -48.8 | -48.8 | -47.6 | -48.8 | -48.7 |
| BQTerrace | -45.9 | -46.5 | -45.3 | -45.8 | -46.4 | -45.2 |
| Summary | | | | | | | | |
| **B** |  |  | -52.0 | -51.5 | -51.6 | -51.9 | -51.5 | -51.5 |

Table 10. BD-rate for the proposed enhanced resolution layer relative to HEVC enhanced resolution single layer anchor including base layer anchor (spatial resolution factor of 1.5)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **B** | 1920x1080 | Kimono | -35.6 | -32.6 | -33.7 | -35.6 | -32.6 | -33.7 |
| ParkScene | -32.9 | -32.2 | -33.5 | -32.8 | -32.2 | -33.4 |
| Cactus | -30.4 | -30.0 | -29.0 | -30.4 | -29.9 | -28.9 |
| BasketballDrive | -26.5 | -27.1 | -26.8 | -26.4 | -27.1 | -26.8 |
| BQTerrace | -26.1 | -26.9 | -26.4 | -26.0 | -26.9 | -26.4 |
| Summary | | | | | | | | |
| **B** |  |  | -30.3 | -29.8 | -29.9 | -30.3 | -29.7 | -29.8 |

### SNR scalability

The proposed enhanced SNR layers were created with a QP change (increasing or decreasing QP value by 1) in the middle of the sequence to match the target bitrate more accurately.

Table 11. BD-rate for the proposed enhanced SNR layer relative to HEVC enhanced SNR single layer anchors

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **A+** | 3840x2048 | Traffic | -35.9 | -31.8 | -29.2 | -35.4 | -31.3 | -28.7 |
| 3840x2160 | PeopleOnStreet | -42.6 | -22.7 | -23.3 | -42.1 | -22.2 | -22.8 |
| **B** | 1920x1080 | Kimono | -38.7 | -30.5 | -27.8 | -38.4 | -30.2 | -27.5 |
| ParkScene | -35.2 | -30.4 | -30.7 | -34.8 | -30.0 | -30.4 |
| Cactus | -35.2 | -27.5 | -19.9 | -34.9 | -27.2 | -19.5 |
| BasketballDrive | -36.0 | -23.1 | -23.7 | -35.8 | -22.8 | -23.4 |
| BQTerrace | -30.9 | -8.7 | 11.3 | -30.7 | -8.5 | 11.7 |
| Summary | | | | | | | | |
| **A+** |  |  | -39.2 | -27.2 | -26.3 | -38.8 | -26.7 | -25.7 |
| **B** |  |  | -35.2 | -24.0 | -18.2 | -34.9 | -23.7 | -17.8 |
| **All** |  |  | -37.2 | -25.6 | -22.2 | -36.8 | -25.2 | -21.8 |

Table 12. BD-rate for the proposed enhanced SNR layer relative to HEVC enhanced SNR single layer anchor including base layer anchor

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Actual bitrate | | | Target bitrate | | |
| Y BD-Rate | U BD-Rate | V BD-Rate | Y BD-Rate | U BD-Rate | V BD-Rate |
| **A+** | 3840x2048 | Traffic | -23.2 | -19.8 | -17.7 | -22.8 | -19.5 | -17.3 |
| 3840x2160 | PeopleOnStreet | -28.5 | -11.9 | -12.2 | -28.3 | -11.6 | -11.9 |
| **B** | 1920x1080 | Kimono | -25.3 | -18.8 | -16.7 | -25.1 | -18.5 | -16.4 |
| ParkScene | -22.8 | -18.8 | -19.3 | -22.6 | -18.6 | -19.0 |
| Cactus | -23.1 | -16.8 | -10.2 | -22.9 | -16.6 | -10.0 |
| BasketballDrive | -23.6 | -12.7 | -13.2 | -23.4 | -12.4 | -13.0 |
| BQTerrace | -21.2 | -1.6 | 15.4 | -21.1 | -1.4 | 15.6 |
| Summary | | | | | | | | |
| **A+** |  |  | -25.9 | -15.9 | -14.9 | -25.5 | -15.5 | -14.6 |
| **B** |  |  | -23.2 | -13.7 | -8.8 | -23.0 | -13.5 | -8.6 |
| **All** |  |  | -24.5 | -14.8 | -11.9 | -24.3 | -14.5 | -11.6 |

## Category 2 (AVC base layer)

### Spatial scalability

### Overall

# Complexity analysis

## Encoding time and measurement methodology

The encoding times were collected from log files generated by the encoder. The encoding times written to the log files were obtained using the same process as used by HM6.1 software.

The following tables show the encoding time for the proposal and HEVC enhanced single layer anchors with the second lowest QP, 22.

Table 13. Encoding time for the proposal for spatial scalability and HEVC enhanced resolution single layer anchor.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Prop. encoding time (sec/QP) | | Anchor encoding time (sec) (QP=22) |
| 2x | 1.5x |
| **A+** | 3840x2048 | Traffic | 36549.27 |  | 26862.69 |
| 3840x2160 | PeopleOnStreet | 46981.98 |  | 37383.33 |
| **B** | 1920x1080 | Kimono | 17111.89 | 19496.05 | 12931.58 |
| ParkScene | 15990.90 | 18132.28 | 11914.82 |
| Cactus | 34831.08 | 39173.64 | 27622.42 |
| BasketballDrive | 39064.19 | 44279.41 | 31274.65 |
| BQTerrace | 41548.20 | 46865.98 | 36344.35 |

Table 14. Encoding time for the proposal for intra-only spatial scalability and HEVC enhanced resolution single layer anchor

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Prop. encoding time (sec/QP) | | Anchor encoding time (sec) (QP=22) |
| 2x | 1.5x |
| **A+** | 3840x2048 | Traffic | 25820.81 |  | 9242.20 |
| 3840x2160 | PeopleOnStreet | 28313.25 |  | 10025.67 |
| **B** | 1920x1080 | Kimono | 9591.78 | 10699.02 | 3560.68 |
| ParkScene | 11227.94 | 11737.35 | 4234.65 |
| Cactus | 22594.50 | 23659.16 | 9099.28 |
| BasketballDrive | 20290.36 | 20736.23 | 8319.67 |
| BQTerrace | 27994.16 | 28967.60 | 11500.36 |

Table 15. Encoding time for the proposal for SNR scalability and HEVC enhanced SNR single layer anchor

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Prop. encoding time (sec/QP) | Anchor encoding time (sec) (QP=22) |
| **A+** | 3840x2048 | Traffic | 55229.29 | 26862.69 |
| 3840x2160 | PeopleOnStreet | 70501.41 | 37383.33 |
| **B** | 1920x1080 | Kimono | 26098.76 | 12931.58 |
| ParkScene | 23932.75 | 11914.82 |
| Cactus | 52379.52 | 27622.42 |
| BasketballDrive | 59346.78 | 31274.65 |
| BQTerrace | 63140.47 | 36344.35 |

## Decoding time and measurement methodology and comparison vs. anchor bitstreams decoded by HM

The decoding times were collected from log files generated by the decoder. The decoding times written to the log files were obtained using the same process as used by HM6.1 software. The I/O times (for output YUV file generation) are included in the reported decoding times.

The following tables show the decoding time for the proposal and HEVC enhanced single layer anchors with the second lowest QP, 22.

Table 16. Decoding time for the proposal for spatial scalability and HEVC enhanced resolution single layer anchor

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Prop. decoding time (sec/QP) | | Anchor decoding time (sec) (QP=22) |
| 2x | 1.5x |
| **A+** | 3840x2048 | Traffic | 93.49 |  | 63.26 |
| 3840x2160 | PeopleOnStreet | 135.56 |  | 94.11 |
| **B** | 1920x1080 | Kimono | 45.21 | 56.64 | 29.19 |
| ParkScene | 43.58 | 52.17 | 30.31 |
| Cactus | 83.40 | 97.28 | 60.15 |
| BasketballDrive | 95.79 | 113.56 | 66.77 |
| BQTerrace | 109.80 | 128.35 | 93.54 |

Table 17. Decoding time for the proposal for intra-only spatial scalability and HEVC enhanced resolution single layer anchor

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Prop. decoding time (sec/QP) | | Anchor decoding time (sec) (QP=22) |
| 2x | 1.5x |
| **A+** | 3840x2048 | Traffic | 187.13 |  | 142.91 |
| 3840x2160 | PeopleOnStreet | 206.46 |  | 150.60 |
| **B** | 1920x1080 | Kimono | 65.51 | 77.43 | 43.18 |
| ParkScene | 80.90 | 92.93 | 62.38 |
| Cactus | 161.43 | 184.65 | 129.48 |
| BasketballDrive | 141.48 | 159.82 | 111.68 |
| BQTerrace | 204.97 | 233.00 | 176.64 |

Table 18. Decoding time for the proposal for SNR scalability and HEVC enhanced SNR single layer anchor

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Resolution | Sequence name | Prop. decoding time (sec/QP) | Anchor decoding time (sec) (QP=22) |
| **A+** | 3840x2048 | Traffic | 128.77 | 63.26 |
| 3840x2160 | PeopleOnStreet | 180.46 | 94.11 |
| **B** | 1920x1080 | Kimono | 61.57 | 29.19 |
| ParkScene | 60.12 | 30.31 |
| Cactus | 114.05 | 60.15 |
| BasketballDrive | 132.87 | 66.77 |
| BQTerrace | 152.86 | 93.54 |

## Description of computing platform used to determine encoding and decoding times reported in sections 5.1 and 5.2

The following machine is used for encoding time measurement for spatial, intra-only spatial scalability an SNR scalability. The encoding was executed on a computer cluster that runs a 64-bit version of Linux. The cluster has 11 machines and 66 cores. All the encoding jobs were launched together and put a in a queue. Each machine ran 5 jobs at a time

|  |  |
| --- | --- |
| **CPU** | Intel i7 990x Extream, 6 Core, 3.46 GHz |
| **Memory** | 6 x 4GB DDR3 SDRAM(1333MHz) |
| **HDD** | 2TB 3.5” SATA-III 7200 rpm, 64MB |
| **OS** | CentOS Linux 64bit |

The following machine is used for decoding time measurement for spatial, intra-only spatial, and SNR scalability. The decoding was executed on a Windows PC running 64-bit windows 7. For decoding time measurements, only one core is occupied by decoder.

|  |  |
| --- | --- |
| **CPU** | Intel i7 990x Extream, 6 Core, 3.46 GHz |
| **Memory** | 6 x 4GB DDR3 SDRAM(1600MHz) |
| **HDD** | 2TB 3.5” SATA-III 7200 rpm, 64MB |
| **OS** | Windows 7 64bit |

## Expected memory usage of encoder

The memory usage of the encoder is the sum of the memory usage of the base layer encoder and the enhancement layer encoder.

The memory usage of the enhancement layer encoder is increased by the following two factors.

* One more picture buffer size (in case of spatial scalability): Since the ILRP is stored for inter-layer prediction at the enhancement layer.
* ALF is enabled.

## Expected memory usage of decoder

Since the proposal employs the multiple loop decoding design, the memory usage of the decoder is the sum of the memory usage of the base layer decoder and the enhancement layer decoder.

The memory usage of the enhancement layer decoder is increased by the following two factors.

* One more picture buffer size (in case of spatial scalability): Since the ILRP is stored for inter-layer prediction at the enhancement layer
* ALF is enabled.

## Complexity characteristics of encoder motion estimation and partitioning selection in enhancement layer(s)

There is no specific difference compared to the HEVC draft spec[2] since they share the similar motion estimation and partitioning selection scheme.

However, the complexity of the encoder motion estimation and partitioning selection is increased by additional reference pictures (i.e. ILRP) as follows.

* The numbers of reference pictures increased by 1: Since the ILRP is inserted into the reference picture lists of the enhancement layer, the numbers of reference pictures are increased by 1.
  + For random access picture: NumOfReference=1 and NumOfReferenceB\_Lx=1
  + For non-random access picture: NumOfReference=5 and NumOfReferenceB\_Lx=3

## Complexity characteristics of decoder motion compensation in enhancement layer(s)

There is no specific difference compared to the HEVC draft spec[2] since they share the similar motion compensation scheme. However, the complexity of the decoder motion compensation is increased because a picture of the reference layer is reconstructed (and upsampled for spatial scalability) for motion compensation.

## Complexity characteristics of encoder intra-frame prediction type and partitioning selection in enhancement layer(s)

There is no specific difference compared to the HEVC draft spec[2] since they share the same intra-frame prediction type and partitioning selection.

## Complexity characteristics of decoder intra-frame prediction operation in enhancement layer(s)

There is no specific difference compared to the HEVC draft spec[2] since they share the same intra prediction operation scheme.

## Complexity characteristics of upsampling filters and transforms specific in enhancement layer(s)

The up-sampling filter is only applied to the texture image of the reference layer as described in Section 2.2. For up-sampling luma and chroma samples, the 8-taps and 4-tpas DCT-IF filters with 6 bits filter coefficients representation are used, respectively.

There is no specific difference compared to the HEVC draft spec[2] since they share the same transform coding scheme.

## Complexity characteristics of quantization and inverse quantization in enhancement layer(s)

There is no specific difference compared to the HEVC draft spec[2] since they share the same scalar dequantization and inverse quantization scheme.

## Complexity characteristics of encoder entropy coding operation in enhancement layer(s)

There is no specific difference compared to the HEVC draft spec[2] since they share the same entropy coding scheme.

## Complexity characteristics of decoder entropy decoding operation in enhancement layer(s)

There is no specific difference compared to the HEVC draft spec[2] since they share the same entropy coding scheme.

## Degree of capability for encoder parallel processing

The encoder parallel processing capabilities for our proposal are similar to those for HEVC draft spec[2].

## Degree of capability for decoder parallel processing

The decoder parallel processing capabilities for our proposal are similar to those for HEVC draft spec[2].

# Software implementation description

The proposed algorithms were implemented using standard C++ programming language.

# References

1. ISO/IEC JTC1/SC29/WG11 and ITU-T SG 16 WP 3, "*Joint Call for Proposals on Scalable Video Coding Extensions of High Efficiency Video Coding(HEVC)*", w12957, 101st MPEG, Stockholm, July. 2012
2. B. Bross, W.-J.Han, J.-R. Ohm, G. J. Sullivan, and T. Wiegand, “High efficiency video coding (HEVC) text specification draft 6”, JCTVC-H1003, 8th JCT-VC meeting, San Jose, Feb., 2012
3. F. Bossen, “Common HM test conditions and software reference configurations”, JCTVC-H1100, 8th JCT-VC meeting, San Jose, Feb., 2012

# Patent rights declaration(s)

**ETRI may have IPR relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**

**Kwangwoon Univ. may have IPR relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**