|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  12th Meeting: Geneva, CH, 14–23 Jan. 2013 | Document: JCTVC-L0042 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Content Adaptive Complexity Reduction Scheme For Quality/Fidelity Scalable HEVC** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or*  *Contact(s):* | 1Hamid Reza Tohidypour, 1,2Mahsa T. Pourazad, and 2Panos Nasiopoulos  2366 Main Mall, Engineering, Vancouver, BC, V6T 1Z4, Canada | Tel: Email: | {htohidyp, pourazad, panosn}@ece.ubc.ca |
| *Source:* | |  | | --- | | 1University of British Columbia (UBC), 2TELUS Communications Inc. | | | |

# Abstract

This contribution suggests a scheme to reduce the complexity of the SNR/Quality/Fidelity scalable HEVC. This scheme utilizes the correlation between the enhancement layers and the base layer to minimize redundant computations while encoding the enhancement layer. The presented scheme adaptively adjusts the motion search range in the enhancement layer based on the motion vector information of the base layer and implements an adaptive early-termination approach for inter and intra prediction mode search in the enhancement layer. The reported results illustrate that this scheme reduces HEVC/SVC’s coding complexity by up to 61.66%, while maintaining the overall bitrate.

**1. Introduction**

One challenge in implementing the scalable extension of HEVC is complexity. In this contribution we propose a scheme to effectively reduce the complexity of the SNR/Quality/Fidelity scalable HEVC. Our scheme utilizes the correlation between the enhancement layers and the base layer to reduce the computational complexity of the inter and intra prediction processes. This is done by utilizing specific coding information from the base layer such as motion homogeneity, prediction modes, and the rate distortion (RD) cost. The rest of this report is organized as follows: Section 2 provides an overview on our proposed method, Section 3 elaborates on the performance evaluation, and Section 4 concludes the report.

**2. our proposed scheme**

**2.1. Adaptive search range adjustment**

In the inter prediction process, selecting a large search range leads to high computational costs, while selecting a small search range produces poor matching results. The optimal motion search range should result in reduced complexity without hampering the compression performance.

In the case of scalable video coding, the MVs of the base layer and those of the enhancement layer are correlated. In our scheme we utilize this correlation to select the proper motion search range for the enhancement layer based on the motion information of the base layer. We first classify the LCUs within each frame in the base layer to different classes in terms of search range as follows:

(1)

where SR indicates the search range defined in the configuration file, MVs represents the motion vectors of the LCU in the base layer , and SRLCU is the adjusted search range of the co-located LCU in the enhancement layer. Note that, depending on the class of the LCU in the base layer, the search range of the co-located LCU in the enhancement layer is adjusted, and all the CUs within that LCU will have the same adjusted motion search range setting. As it can be observed from (1), the search range of the enhancement layer LCU can become quite small, depending on the LCU class, which can significantly reduce the overall computational cost.

**2.2. Early termination mode search**

The objective here is to reduce the computational complexity by implementing an early termination (ET) mode-search at the enhancement layer, so that the encoder does not need to search all the modes.

The HEVC encoder, in the inter/intra prediction mode selection process, calculates the Rate Distortion (RD) cost for each mode and the one with minimum RDcost is selected. In mode search, if the RD cost of the current to-be-coded CU in the enhancement layer is predicted from the already coded CUs, the mode-search can be terminated before going through all the modes and thus the overall computational complexity is significantly reduced. In order to find a prediction for the RD cost of the current CU in the enhancement layer, the RD cost of the already coded CUs in the enhancement layer and their co-located CUs in the base layer is utilized. Figure 1 shows an example of the arrangement of the CUs whose information is utilized to predict the RD cost of the to-be-coded CU in the enhancement layer. The neighboring CUs in the enhancement layer are similar to the candidates that HEVC chooses for the merge mode motion search. Inspired by [4], we assume that there is an additive model between the RD cost of the CUs in the enhancement layer and their co-located CUs in the base layer as follows:

(2)

where RDcostECpredict is the predicted RD cost of current CU in the enhancement layer, *RDcostBc* is RD cost of the co-located CU in the base layer, *RDcostET*, *RDcostEL, RDcostETL*and *RDcostETR*denote the RD cost of the four spatial neighbors of the current CU (see Figure 1 for the arrangement of CUs), *RDcostBT*, *RDcostBL*, *RDcostBTL and RDcostBTR* are the RD cost values of the corresponding CUs in base layer, and *α*0, *α*1,*α*2 and *α*3 are weighting constants, which will be computed in the following sub-section.

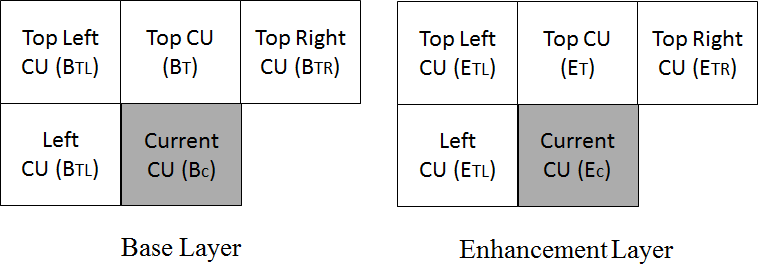


Figure 1 Current CU and its four spatial neighbors of base layer and Enhancement layer

Once the predicted RD cost for the current CU is available, a threshold for early termination of mode search in the enhancement layer is defined as follows:



Figure 2 Block diagram of our HEVC-based complexity reduction scheme

*Thr = min*(RDcostET, RDcostEL, RDcostETL, RDcostETR, RDcostECpredict) (3)

Using this threshold the encoder terminates the mode search if the RD cost of a mode is less than the threshold, and selects that mode as the best one. Otherwise, it continues testing other modes until this criterion is met. Note that this scheme is applied to the CUs with at least two already-coded neighboring CUs. Figure 2 shows the flowchart of the proposed scheme.

In the case where the size of co-located CUs in the base layer is not similar to the one in the enhancement layer, the RD cost of the co-located CU is normalized to its size and the RD cost used in our calculation is updated as follows:

(4)

where W and H are respectively the width and height of the co-located CU in the base layer, w and h is the width and height of the current CU in the enhancement layer, λ and B are the Lagrangian constant value and the bit-cost of the co-located CU in the base layer respectively, and *RDcostn* is the RD cost value to be used in finding the predicted RD cost and the threshold.

**2.3. Determining weighting constants**

In order to find the proper weighting constants in equation (2), the Linear Least Square method is used. Our objective is to minimize the difference between the predicted RD cost and the real RD cost of the best mode (without using ET) for the current to-be-coded CU in the enhancement layer. This objective is formulated as follows:

(5)

where *S* is a matrix that contains the real RD cost values of the best modes selected by HEVC for the current CU in the enhancement layer (RDcostEC) divided by RDcostBC, *S′* denotes a matrix which contains the predicted RD cost of the current CU (RDcostECpredict) divided by RDcostBC. We can re-write *S′* as follows:

(6)

where

i=1,2,3,…,n

Thus, the weighting constants are calculated as follows:

(7)

We use a train dataset (five representative video sequences) to calculate the weighting constants. We code the video streams, record the real RD cost values, calculate the predicted RD cost based on equation (5), and find the weighting constants based on equation (7). In the case that all four spatial neighbors (T, L, TL and TR) are available (see Figure 1), the estimated weighting constants are as follows: [α0, α1, α2, α3] = [0.33, 0.32, 0.18, 0.17]. When RDcostETR is not available, [α0, α1, α2, α3] = [0.4027, 0.3913, 0.2059, 0]. If RDcostETL is not available – which means that the RDcostEL is not available either - we use two upper neighbors to predict the RD cost, and the weighting constants are [α0, α1, α2, α3] = [0.5154, 0, 0, 0.4846]. The weighting constants of the top and left neighboring CUs when available are larger than the others, denoting that they are more correlated with the current CU.

**3. results and discussion**

In our experiment, four test videos from the data set provided by MPEG for the HEVC Call for proposals [5] were used (see Table I). Note that the train data used for finding the weighting constants was not included in our test videos. Our method was implemented on the HEVC software (HM 8.1), with Random Access High Efficiency (RA-HE) configuration (hierarchal B pictures, GOP length 8, SAO and RDOQ were enabled) [6]. The QPs used for the base layer and enhancement layer (QpB , QpE) are as follows: (22, 26), (32, 28), (38, 32) and (40, 36).

Figure 3 shows the RD curves of the test video sequences, where bitrate is the average total bitrate of the scalable video stream (base layer+ enhancement layer). As it can be observed, the proposed scheme minimally hampers the bitrate.

Figure 3 also illustrates the percentage of mode-search complexity reduction for each stream. In our study the complexity is computed based on the number of times the encoder searches for the best mode. For example, for inter prediction, for every search point the complexity measure is equal to 1. For the skip mode, the complexity measure is equal to 1. For the Merge mode the complexity is up to 5, depending on the available candidates. By adding up these complexity values when coding the enhancement layer, we find the total complexity measure. For intra modes, we also compute the number of candidates which the encoder checks to encode a CU. The complexity curves in Figure 3 indicate that the proposed scheme substantially reduces the computational cost without hampering the total bitrate.

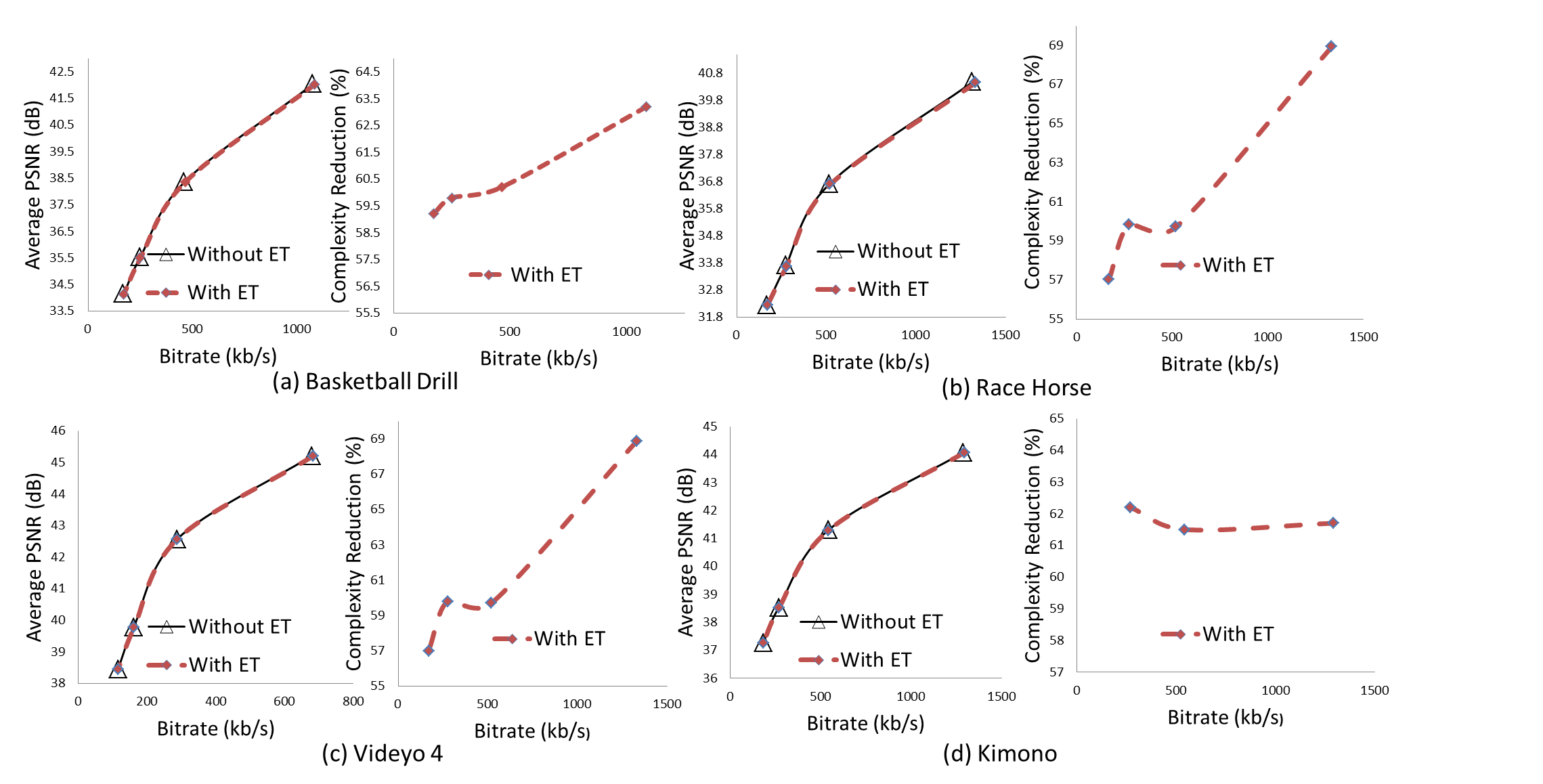


Figure 3 Rate Distortion curves and the complexity reduction curves of the proposed scheme.

Table 1 summarizes the effect of our scheme in terms of bitrate, PSNR and complexity for each stream. As it can be observed, the proposed scheme reduces the complexity by up to 61.35% at the cost of maximum 1.499% bitrate increase.

Table I Impact of the proposed scheme on Bitrate, PSNR and Complexity

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Name | Resolution, Frame Rate (fps) | Average PSNR Degrade | Average Bitrate Increase | Average Complexity Reduction |
| BasketballDrill | 832x480, 50 | 0.037dB | 0.873% | 59.2% |
| Kimono | 1920x1080, 24 | 0.059 dB | 1.499% | 61.35% |
| Race Horse | 832x480, 30 | 0.014 dB | 0.535% | 60.67% |
| Videyo 4 | 1280x720, 60 | 0.016 dB | 0.515 % | 60.67% |

**4. conclusions**

We proposed a content adaptive complexity reduction scheme for SNR/Quality scalable HEVC. In the proposed scheme the information of the base layer is utilized to facilitate the inter prediction and intra prediction mode selection process in the enhancement layer by avoiding redundant computations.

Performance evaluations show that the proposed scheme reduces the complexity by up to 61.35% at the cost of maximum 1.499% bitrate increase.

**5. References**

[1] A. H. Schwarz, D. Marpe, T. Wiegand, “Overview of the Scalable Video Coding Extension of the H.264/AVC Standard”, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17 (9), pp. 1103-1120, Sep. 2007.

[2] J.R. Ohm, “Scalable Video Coding” (Presentation Slides), Shenzhen workshop, Oct. 2007.

[3] M. T. Pourazad, C. Doutre, M. Azimi, and P. Nasiopoulos, “HEVC: The New Gold Standard for Video Compression,” IEEE Consumer Electronic Magazine, vol.1 , issue 3, pp. 36-46, July 2012.

[4] L. Shen, Z. Zhang, “Content adaptive motion estimation algorithm for coarse-grain scalability”, IEEE Trans. on Image Processing, pp. 2582-2591, Vol. 21, No.2, March 2012.

[5] ISO/IEC JTC1/SC29/WG11,“Joint Call for Proposals on Video Compression Technology”, N11113, January 2010.

[6] F. Bossen, “Common test conditions and software reference configurations”, JCTVC-F900, July 2011.

**6. PATENT RIGHTS DECLARATION**

TELUS/UBC may have current or pending patent rights 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).