Adaptive Interpolation Filter Based on Deblocking Filter Characteristics for HEVC

Yunseok Song, Jung-Ah Choi, and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST), Gwangju, Rep. of Korea
{ysong, jachoi, hoyo}@gist.ac.kr

***Abstract*—This paper proposes an adaptive interpolation filter for the newest video coding standard, high efficiency video coding (HEVC). For motion estimation and compensation, 8-tap and 4-tap filters are used for luma and chroma samples, respectively. The proposed filter employs boundary strength values which are computed by the deblocking filter. Prior to interpolation, based on the boundary strength value, each region is classified into either complex or simple region. This is applied to luma interpolation only. The proposed method is implemented on HEVC test model (HM) 11.0. Simulation results indicate 1.35% average gain by the proposed method compared to the conventional HEVC.**

Keywords—HEVC; interpolation; deblocking filter

I. Introduction

High efficiency video coding (HEVC) allows state-of-the-art video compression. This standard was developed by the joint collaborative team on video coding (JCT-VC) formed by moving picture experts group (MPEG) of ISO/IEC and video coding experts group (VCEG) of ITU-T. Compared to the previous video coding standard, H.264/AVC, HEVC enhanced conventional coding tools such as intra/inter prediction, transform/quantization, entropy coding and deblocking filter [1, 2]. In particular, the number of intra prediction modes is increased to 35. Further, symmetric as well as asymmetric partitions are considered in inter prediction. For entropy coding, context based adaptive arithmetic coding (CABAC) is solely used. Sample adaptive offset (SAO) is a newly introduced tool used in loop filtering. This tool is used for pixel-wise error compensation for artifact reduction. In addition to such techniques, flexible prediction units with varying sizes, i.e., from 4×4 to 64×64 are used.

In motion estimation and compensation, interpolation is is performed to enable quarter-sample motion vector accuracy. For luma and chroma samples, 8-tap and 4-tap filters are used, respectively. The filter coefficients are derived from discrete cosine transform (DCT). However, the DCT-based interpolation filter (DCT-IF) does not consider video characteristics such as the complexness of the region. For more complex region, a longer tap interpolation filter is more effective.

Several adaptive interpolation filtering methods have been proposed. In [3], multiple sets of filter coefficients are generated to optimize each region. In [3], filter coefficients for vertical and horizontal directions are separately derived using minimization of prediction error.

In this paper, we propose an adaptive interpolation filter which uses boundary strength (Bs) values. Such values are computed by the deblocking filter as part of the in-loop filter. Depending on the sum of Bs values, each region is classified as complex or simple. We use a 12-tap filter for complex regions and an 8-tap filter for simple regions. This is applied to luma sample interpolation only since region characteristics are more related to luma sample interpolation.

II. Deblocking Filter

In HEVC, loop filtering stage consists of deblocking filter and SAO. These tools compensate errors to enhance the overall picture quality prior to the outputting process. The deblocking filter is executed first, followed by SAO. Such a tool is designed to reduce blocking artifacts which exhibit sudden variation of pixels at block boundaries [4]. They are caused by block-based transform coding followed by quantization. Prediction of adjacent blocks could also cause this problem.

Depending on the Bs value and other thresholds, one of the three actions is taken: no filtering, normal filtering or strong filtering. Bs is determined by several factors. Filtering is applied only when the block boundary is either a prediction unit (PU) boundary or a transform unit (TU) boundary.

A. Boundary Strength Decision Process

For calculation of Bs, 8×8 blocks are used. First, if at least one of the blocks is intra coded, Bs is the highest value, two. In the case that neither is intra coded, Bs is one if one of the following conditions are satisfied.

1) One of the blocks has non-zero coded residual coefficients.

2) The blocks have different reference pictures.

3) They have different number of motion vectors.

4) The absolute difference of their motion vectors (MV) is greater than four.

The Bs value is zero if none of such conditions are satisfied. This means the block boundary is smooth with little variation. Thus, no filtering is necessary in this case. Fig. 1 represents the Bs decision process.

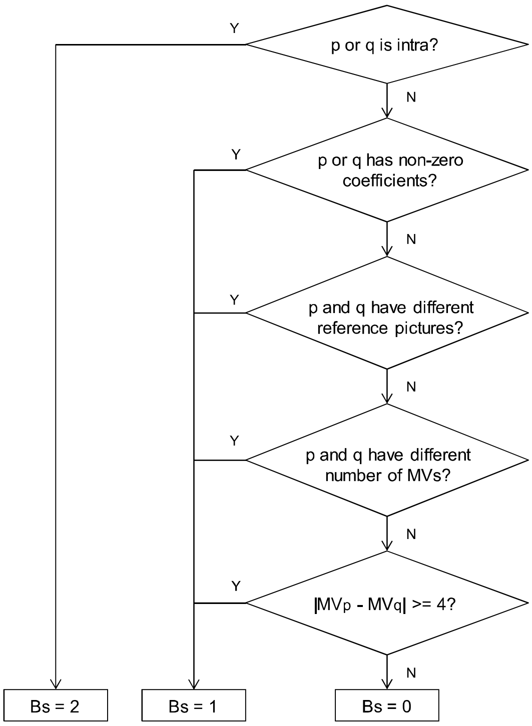


Fig. 1. Block strength decision process

B. Filtering

Deblocking process is executed if Bs is greater than zero for luma samples, and one for chroma samples. Deblocking is performed by either normal filtering or strong filtering. The type of filtering is determined by several conditions based on sample values and two defined thresholds. The specifics related to these are not covered in this paper since the Bs derivation is the main interest.

III. PROPOSED METHOD

The proposed method first stores sum of Bs values for each CU. Prior to motion estimation of the successive frame, the Bs sum is compared with the average Bs sum of the entire frame; if greater, the block is determined as complex region, otherwise, simple region. A 12-tap DCT-IF is applied to complex regions. For simple regions, the conventional 8-tap filter is used. Fig. 2 represents the flowchart of the proposed method.

A. Boundary Strength Usage

Bs values indicate the complexness of the region. In order to use this information, we add Bs values for each CU during the deblocking filtering process. Since deblocking is applied in both horizontal and vertical directions, Bs values for both directions are added. High Bs sum implies numerous Bs values possessing value two, i.e., CU includes many intra blocks, thus, spatial details are present. On the contrary, there will be many Bs values possessing zero for low Bs case, indicating smooth transition.

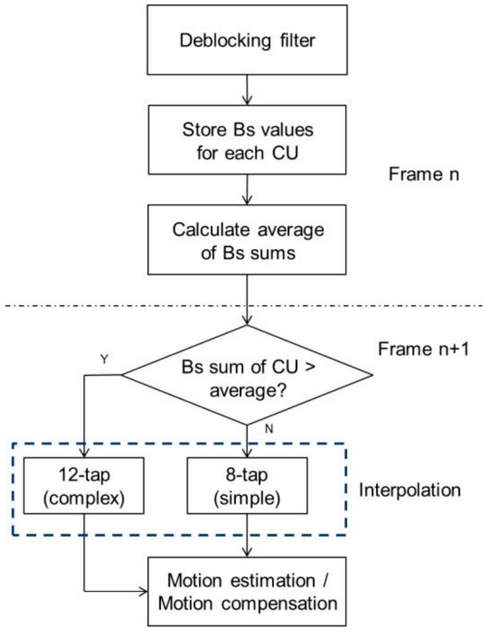


Fig. 2. Flowchart of the proposed method

B. Interpolation Filter

In the early stage of HEVC standardization, 12-tap DCT-IF was used [7]. Despite its better quality in general, complexity concerns were raised. After various evaluations, 8-tap DCT-IF was adopted.

When applying the interpolation filter, horizontal 1D filter is applied first and then vertical interpolation is executed. The same interpolation filter is used for both directions in this paper. Since vertical and horizontal Bs values exist, our further research aims to exploit these and separate the interpolation filter for increased accuracy.

TABLE I. INTERPOLATION FILTER COEFFICIENTS (LUMA)

Filter type	Fractional Position	Filter Coefficients
8-tap	1/2	{-1, 4, -11, 40, 40, -11, 4, 0}
	1/4	{-1, 4, -10, 58, 17, -5, 1, 0}
12-tap (precision=8)	1/2	{-1, 8, -16, 24, -48, 161, 161, -48, 24, -16, 8, -1}
	1/4	{-1, 6, -12, 21, -43, 229, 75, -30, 17, -10, 5, -1}
12-tap (precision=6)	1/2	{-1, 3, -4, 6, -12, 40, 40, -12, 6, -4, 3, -1}
	1/4	{-1, 2, -3, 6, -11, 58, 19, -8, 4, -3, 2, -1}

IV. SIMULATION RESULTS

The proposed method is implemented on HEVC test model (HM) 11.0. Four sequences with all different resolutions were tested: BQSquare (416×240), PartyScene (832×480), Johnny

(1280×720) and BQTerrace (1920×1080). The number of coded frames is 100. Fig. 3 shows the test sequences. Common test conditions in [5] were used under the main profile with the random access mode. Quantization parameters (QP) are set to 22, 27, 32 and 37. Table II and Table III represent bitrate and PSNR results of HM 11.0 and the proposed method, respectively. Table IV reports the Bjontegaard delta rate (BD-rate) [6] for each sequence and the average.

High gains were achieved in BQSquare and PartyScene simulations. Such sequences possess a high portion of textured regions. On the other hand, losses are observed in Johnny and BQTerrace simulations. This is due to their high resolutions and relatively less textured areas.

TABLE II. CODING PERFORMANCE (HM 11.0)

Sequence	QP	Bitrate (kbps)	Y-PSNR (dB)
BQSquare (416×240)	22	1811.72	38.64
	27	816.47	35.04
	32	420.48	32.06
	37	235.22	29.26
PartyScene (832×480)	22	8836.83	38.13
	27	4228.74	34.41
	32	2073.72	31.13
	37	1022.65	28.15
Johnny (1280×720)	22	2209.18	42.86
	27	848.88	41.40
	32	435.61	39.60
	37	244.82	37.31
BQTerrace (1920×1080)	22	39800.43	37.99
	27	10005.10	35.55
	32	3652.04	33.79
	37	1673.24	31.76

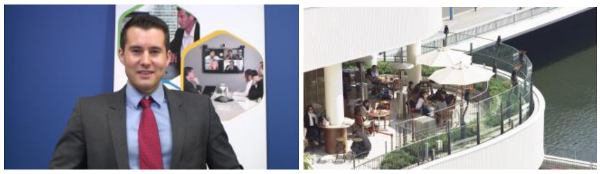
TABLE III. CODING PERFORMANCE (PROPOSED METHOD)

Sequence	QP	Bitrate (kbps)	Y-PSNR (dB)
BQSquare (416×240)	22	1756.21	38.69
	27	784.64	35.19
	32	413.39	32.25
	37	232.88	29.38
PartyScene (832×480)	22	8650.50	38.18
	27	4164.30	34.51
	32	2055.52	31.22
	37	1021.14	28.20
Johnny (1280×720)	22	2217.54	42.84
	27	851.36	41.38
	32	436.58	39.59
	37	246.19	37.32
BQTerrace (1920×1080)	22	40369.82	37.96
	27	10059.49	35.50
	32	3664.05	33.76
	37	1675.78	31.75

TABLE IV. PERFORMANCE EVALUATION (BD-RATE)

Sequence	BD-rate (%)
BQSquare	−5.65
PartyScene	−3.02
Johnny	0.85
BQTerrace	2.43
Average: −1.35%	

(a) BQSquare (b) PartyScene



(c) Johnny (d) BQTerrace

Fig. 3. Test sequences

V. CONCLUSION

In this paper, we proposed an adaptive interpolation filter for HEVC. The conventional method uses only the 8-tap DCT-IF for luma sample interpolation. Since HEVC does not consider input video characteristics in motion estimation and compensation, we employ an adaptive interpolation filter. The proposed method stores the sum of Bs values for each CU. The average of Bs sums in the entire frame is them used as a threshold for the next frame. Prior to motion estimation, if the Bs sum is greater than threshold, a 12-tap DCT interpolation is performed since the region is classified as complex region. Otherwise, the conventional method is executed. Simulation results report 1.35% gain on average. The proposed method performed better on sequences with many textured regions which means the 12-tap filter is used more.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (No. 2013-067321)).

REFERENCES

[1] B. Bross, W.J. Han, J.R.Ohm, G.J. Sullivan, Y.K. Wang, and T. Wiegand, “High Efficiency Video Coding (HEVC) Text Specification Draft 10,” JCTVC-L1003, Jan. 2013.

[2] S. Matsuo, S. Takamura, and H. Jozawa, “LCU-based Adaptive Interpolation Filter,” in Proc. Picture Coding Symposium (PCS) 2012, pp. 393-396, May, 2012.

[3] S. Wittmann and T. Wedi, “Separable Adaptive Interpolation Filter for Video Coding,” in Proc. International Conference of Image Processing (ICIP) 2008, pp. 2500-2503, Oct. 2008.

[4] A. Norkin, G. Bjontegaard, A. Fuldseth, M. Narroschke, M. Ikeda, K. Andersson, M. Zhou, and G. Van der Auwera, “HEVC Deblocking Filter,” IEEE Trans. Circuits and Systems for Video Technology, vol. 22, no. 12, pp. 1746-1754, Dec. 2012.

[5] F. Bossen, “Common Test Conditions and Software Reference Configurations,” JCTVC-L1100, Jan. 2013.

[6] G. Bjontegaard, “Calculation of Average PSNR Differences Between RD-curves,” VCEG-M33, April 2001.