Efficient Depth Map Coding Algorithm for 3D Video System using HEVC

Da-Hyun Yoon, Jae-Il Jung, and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro, Buk-gu,
Gwangju 500-712, Republic of Korea
{yoon, jijung, hoyo}@gist.ac.kr

***Abstract*—In this paper, we propose algorithms for the efficient depth map sequence coding in the high efficiency video coding (HEVC). The proposed algorithm is designed to prevent the inefficiency during inloop filter process and implement the intra prediction precisely. The unique characteristic of depth map sequence is that most regions have smooth depth values and abrupt changes are occurred in boundary regions. Moreover, In HM 3.0, 16 filter coefficient sets are utilized to compensate the error. However, 16 coefficients sets can be overhead when a coefficient set is enough to depth map sequence coding efficiently. We utilize these characteristics to improve the conventional inloop structure and the intra prediction method in the HEVC. In order to evaluate the proposed algorithm, we have implemented the proposed algorithm into the HEVC test model (HM) 3.0. Experimental results show that our proposed algorithm provide higher coding performance than HM 3.0. In terms of BDBR and BDPSNR, our algorithms reduce bitrate and increase PSNR by 0.4% and 0.02 dB, respectively.**

Keywords*: *HEVC, Depth map, Intra prediction, Inloop filter, Rate distortion optimization

I. INTRODUCTION

Recent advantages in video capturing and display technologies will further increase the presence of high and ultra-high definition video contents in multimedia mass market applications. The compression capabilities of the state of art H.264/AVC video coding standard must be improved to accommodate the higher compression efficiency required by these applications. This goal is currently gaining evidence with the standardization activities in the HEVC project. These activities are the result of a successful call for proposals (CfP) [1], issued in January 2010, by joint collaborative team on video coding (JCTVC).

The perceived quality is improved by means of two kinds of inloop filters which reduce the quantization artifacts in HEVC. These two kinds of inloop filters are the adaptive H.264/AVC deblocking filter [2] and the symmetric Wiener filter [3]. The H.264/AVC deblocking filter was designed to filter the blocking artifacts while preserving image edges. Furthermore, this deblocking filter allows modulating the amount of filtering for each block edge by means of two offsets. This modulation may be performed by means of an objective quality metric able to express the subjective impact of the quantization blocking artifacts.

In this paper, we introduce intra prediction and inloop filter structure for depth map sequence. Since the properties of depth map sequence coding is different from the texture sequence coding and the HEVC has been developed for texture sequence, we have large room to improve the performance of the HEVC. After referring to the intra prediction and inloop filter in the HEVC, we explain the properties of depth sequence. Then, we propose the algorithm by using the properties. Finally, we show the experimental results and conclude this paper.

II. APPROACH OF HEVC VIDEO CODING

The HEVC is currently developing the next generation video coding standard referred to as the HEVC. The HEVC is expected to provide around 50% improvement in coding efficiency compared to H.264/AVC. Furthermore, the HEVC is intended for high resolutions and higher frame rates. To attain this efficiency, lots of techniques were proposed. Among the techniques, the intra prediction and inloop filter has been developed largely.

A. Intra Prediction

In order to improve coding efficiency, amount of techniques are proposed in the HEVC. Above all, The intra prediction has been developed largely compared to other coding techniques [4]. Planar prediction and angular prediction are adopted in the HEVC and implemented in HM 3.0.

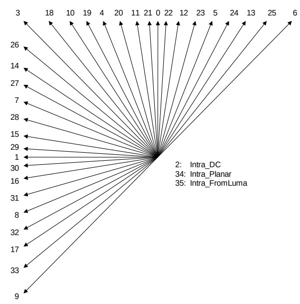


Figure 1. Intra prediction mode direction

Angular prediction is especially efficient when coding directional block. So as to predict various directions, HM 4.0 provides direction up to 34 as shown in Fig. 1. If predictive direction is perpendicular to horizontal direction, predictive direction is determined with disparity between the bottom row of the block and reference row above. Otherwise, predictive direction is determined with displacement between the rightmost column of the block and reference column left. In such case, displacement is set as pixel unit. The predictive pixel is interpolated according to distance from the reference pixel. Fig. 2 shows 6th row having +1 pixel displacement prediction process. The projections of the pixels now fall at the 6/8 sub pixel location (circles) in the reference row of pixels.

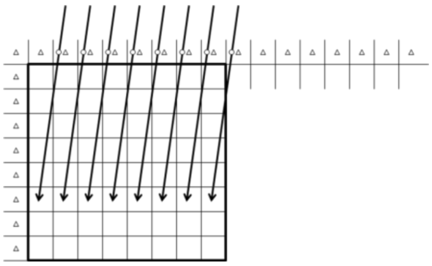


Figure 2. Example of angular prediction

B. Deblocking Filter

In the HEVC, a deblocking filter is adopted to remove the blocking artifact around the CU. A boundary strength (Bs) parameter is assigned to every boundary between two CU to preserve details and remove blocking artifacts. If filtering conditions are satisfied, deblocking filter process will be applied. If boundary strength is equal to 0, filtering process does not take place. Otherwise, filtering on the edge takes place. Fig. 3 shows that how to decide Bs and control offset depending on condition. Moreover, Bs of chrominance is inherited from that of luminance value.

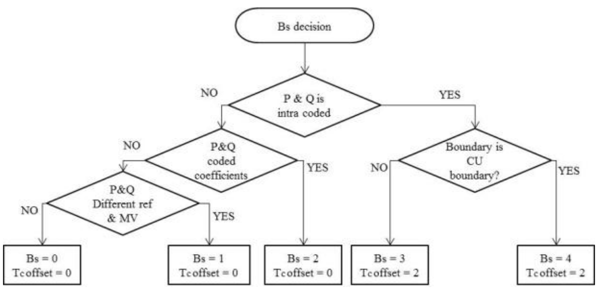


Figure 3. Condition of Bs and control offset

C. Adaptive Loop Filter

Basic concept of adaptive loop filter (ALF) is adaptive filtering of reconstructed signal after finishing sample adaptive offset (SAO) and deblocking filter by using the filter coefficients given in the slice header. All ALF information such as filter coefficients, ALF on-off information, and filter shape is signalized via slice header. For luma ALF processing, ALF on-off signaling is performed in each CU level. Filter coefficients for each coefficient sets are derived based on the Wiener filter design. Filter coefficients for each class are derived by

$$\sum_{i=0}^{N-1} W_i A_{i,j} = C_{j,0} \quad (1)$$

A cross-correlation matrix between the original and reconstructed pixel and auto auto-correlation matrix of the reconstructed pixels are utilized to generate filter coefficients. All CU are marked as on-off flag. The control map consists of on-off flags for CU. The following two SSE of each CU are calculated by

$$SSErecon = (orig(i,j) - recon(i,j))^2 \quad (2)$$

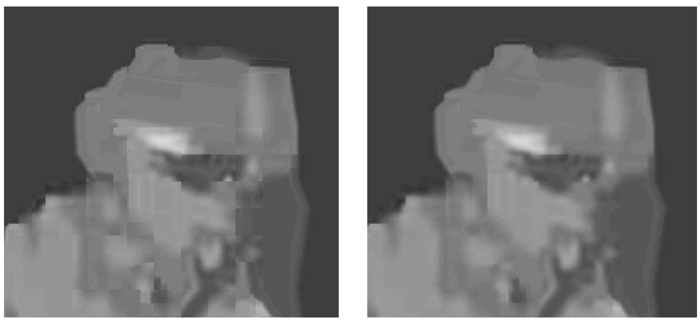
$$SSEalf = (orig(i,j) - filter(i,j))^2 \quad (3)$$

Where $orig(i,j)$, $recon(i,j)$ and $filter(i,j)$ correspond to the original pixels, the reconstructed pixels, and the filtered pixels. If *SSErecon* is smaller than *SSErecon*, the CU is marked as filtered and on-off flag for CU is set to one. Otherwise, the block is marked as not filtered and on-off flag is set to zero.

III. Efficient Depth Map Coding Algorithm

A. Characteristics of Depth Map Sequences in HEVC

In depth video, most regions in picture are smooth and homogeneous except for the boundary regions. The boundary regions have abrupt changes around the edges. Especially, deblocking filter could modify the boundary regions and induce additional bit. Although deblocking filter have the advantage for subjective quality, it is inefficient to use deblocking filter with respect to RD cost as shown in Fig. 4.



(a) (b)

Figure 4. Result of deblocking filtering:
(a) Before deblocking filtering, (b) After deblocking filtering

HM provides ALF filter to improve sequence quality. After deblocking process, ALF is adapted in the reconstruction frame. Each process is performed in CU. If using ALF have lower RD cost than non-ALF CU, HM encode flag bit and coefficients information. Each CU consists of quad tree-based structure and performs optimization, respectively. If ALF filter flag is on, HM encode 16 coefficient sets as exponential golomb code. However, 16 coefficient sets can be overhead when a coefficient set is enough to depth video coding efficiently. Fig. 5 shows that most of slices represent better efficiency when

ALF filter adopt one coefficient set or the frame combining non passing and passing deblocking filter.

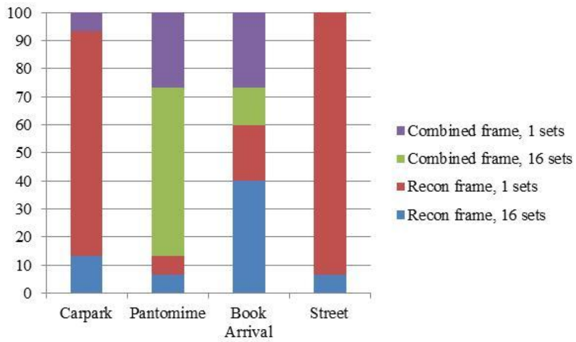


Figure 5. Proposed intra prediction algorithm

Compared to texture sequence, depth sequences have low variable values. However, objects have not gradient change, because each object has same distance from camera. Although there are various angular modes in the intra prediction in the HEVC, this modes is optimized into texture sequences and inefficient to depth map sequences. Since planar mode is selected largely as the best mode in depth map sequence, we can specify planar prediction according to direction such as horizontal and vertical.

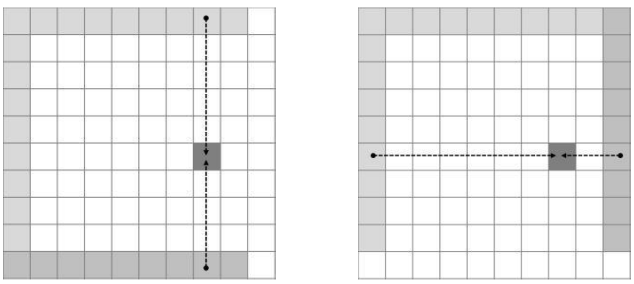
B. Proposed Intra Prediction

Proposed Intra Prediction using directional planar prediction method is as illustrated in Fig. 6. In case of vertical planar prediction, each of kth bottom row is filled with subtraction between bottom left and kth top row and lth row and kth column is calculated by

$$Pixel[l \times stride+k] = Top\_Row[k] + Bottom\_Row[k] \times l/Block\_Size. \quad (4)$$

In case of horizontal planar prediction, each of lth right column is filled with subtraction between top right most and lth left column and lth row and kth column is calculated by

$$Pixel[l \times stride+k] = Left\_Column[l] + Right\_Column[k] \times k/Block\_Size. \quad (5)$$



(a) (b)

Figure 6. Proposed intra prediction algorithm
(a) Vertical planar prediction, (b) Horizontal planar prediction

C. New Structure of Inloop Filter

In HM software, the deblocking filter and adaptive loop filter is connected in order to filter the reconstructed frame, and the two filters are designed independently. Thus, some pixels can be filtered twice. The over filtering can degrade the performance of both filters, thus this structure should be modified to achieve the best performance as possible.

HM adopt 16 pass encoding scheme, this process include 3 iterations of redesign filter process in each depth. Consequently, there are 16 coefficient sets are generated to compensate error between original and reconstructed frames. The properties of depth video is different from texture sequence, 16 coefficient sets could be inefficient when the cost of 1 coefficient is smaller than that of 16 coefficient sets.

We propose new structure of inloop filter by classifying the frame into four categories. In order to avoid over filtering, the proposed structure uses the frame which modified by deblocking filter and untouched by deblocking filter. Let us denote the original input frame as *O*, the reconstructed frame before deblocking filter as *S*, and the reconstruction after deblocking filter as *s'*, frame generated by weighted sum between before and after deblocking filter as *s''*. The weighting factor is determined as 0.9 if quantization parameter (QP) is less than 25. Otherwise, the weighting factor should be 0.5 [6]. There is relationship between weighting factor and QP. We denote that *Category 1* and *Category 3* use 16 pass encoding scheme, and *Category 2* and *Category 4* use one pass encoding scheme. The classification is done by comparing the cost of each category. The result frame having smallest cost is used as best method with respect to RD cost. The structure is illustrated in Fig. 7.

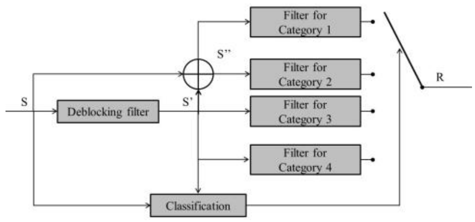


Figure 7. Proposed inloop filter structure

IV. Experimental Results

A. Experimental Condition

The proposed algorithms, new structure of inloop filter and planar prediction optimized to depth video coding, are implemented into the HM 3.0 reference software of the HEVC [6] to compare the performance of the proposed algorihthm. Test sequence and experimental conditions are described in Table 1. We used high efficiency (HE) condition [7] decided in the HEVC standard.We utilize high resolution depth sequences to show better performance as least 1024×768 resolution. Goal of the HEVC is to show high efficiency codec in high resolution and super high resolution. Multiview standard in MPEG is trying to use the HEVC with H.264/AVC. As the size of the depth video sequences is bigger, the necessity of the HEVC increases.

TABLE I. EXPERIMENTAL CONDITION

Parameter	Value	Description
FrameToBeEncoded	50	Number of coded frames
IntraPeriod	1	Period of I-Frame
GOPSize	1	GOP size
QP	22, 27, 32, 37	Quantization parameter
MaxCUWidth	64	Maximum coding unit width
MaxCUHeight	64	Maximum coding unit height
MaxPartitionDepth	4	Maximum coding unit depth
SymbolMode	1	0:LCEC, 1:CABAC
LoopFilterDisable	0	Disable loop filter
ALF	1	Adaptive loop filter

B. Experimental Results

Table 2 shows that experimental results of the proposed algorithm. We reduce Bjontegaard delta bitrate (BDBR) and PSNR (BDPSNR) by 0.4% and 0.02 dB, respectively. Especially, we can save bitrate up to 0.7% in terms of BDBR in Pantomime sequence. The reason why Pantomime sequence represents the best performance is that lots of boundary region which changes gradually is selected as proposed planar prediction. Most of gain comes from the proposed planar prediction method.

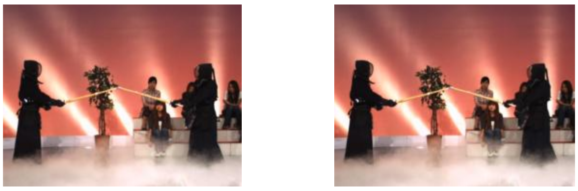
TABLE II. EXPERIMENTAL RESULTS

Sequence	BDBR	BDPSNR
Book Arrival	-0.48	0.026
News	-0.65	0.04
Balloons	-0.33	0.02
Café	-0.25	0.017
Kendo	-0.39	0.021
Pantomime	-0.7	0.03
Carpark	-0.38	0.025
Hall	-0.01	0.00
Street	-0.41	0.013
Average	-0.40	0.02

TABLE III. EXPERIMENTAL RESULTS FOR PROPOSED ALGORITHM

Sequence	QP	HM 3.0	Proposed algorithm	ΔPSNR
Book Arrival	22	34.0757	34.0856	0.010
	27	34.0629	34.0658	0.003
	32	33.9476	33.9505	0.003
	37	33.7679	33.7806	0.013
Newspaper	22	32.2934	32.3025	0.009
	27	32.3203	32.3257	0.009
	32	32.3212	32.3289	0.005
	37	32.3146	32.3269	0.012
Kendo	22	359165	35.9194	0.003
	27	36.0252	36.0216	0.004
	32	36.1117	36.1216	0.010
	37	36.1785	36.1790	0.001

In order to synthesize the intermediate virtual view, we used the 4th and 6th views in Newspaper, the 6th and 8th views in Book Arrival, and the 1th and 3th views in Kendo. The results of PSNR differences between the synthesized sequence and original sequences are shown in Table 3. VSRS rendering software was used for this task [8]. The proposed algorithm slightly increases rendering quality. Fig. 8 shows that there is no significant rendering quality degradation between the HM 3.0 and the proposed algorithm.



(a) (b)

Figure 8. Synthesized result of Kendo sequence
(a) HM 3.0, (b) Proposed algorithm

V. CONCLUSION

In this paper, we propose two efficient depth map sequence coding algorithms using the high efficiency video coding (HEVC). Since the property of depth map sequence coding is different from the texture sequence coding, and the HEVC has been developed for texture sequence, we have large room to improve the performance of the HEVC. The first algorithm was especially designed to improve efficiency of boundary regions modified by deblocking filter and reduce the overhead induced from adaptive loop filter (ALF). The second algorithm optimizes the intra prediction for depth map sequence coding by using the property of smooth depth level change over a frame. The experimental results show that our proposed algorithm provides higher coding performances than the intra prediction and inloop filter method of the HEVC with respect to BDBR and BDPSNR as well as quality of synthesized intermediate views. The proposed algorithm reduces bitrate by 0.4% and increase PSNR 0.02 dB under equivalent PSNR values and bitrates, respectively.

ACKNOWLEDGMENT

This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2011.

REFERENCES

[1] ISO/IEC JTC1/SC29/WG11, “Joint Call for Proposal on Video Compression Technology,” N11113, Jan. 2010.

[2] P. List, A. Joch, J. Lainema, G. Bjøntegaard and M. Karczewicz, “Adaptive deblocking filter”, IEEE Trans. on Circuits and Syst. for Video Technol., vol. 13, no. 7, pp. 614-619, Jul. 2003.

[3] T. Chujoh, N. Wada and G. Yasuda, “Inloop filter using block-based filter control for video coding,” International Conference on Image Processing (ICIP), pp. 1013-1016, Nov. 2009.

[4] ISO/IEC JTC1/SC29/WG11, “Unification of the Directional Intra Prediction Methods in TMuC,” JCTVC-B100, July 2010.

[5] Q. Chen, Y. Zheng, P. Yin, X. Lu, J. Sole, “Classified Quadtree-based Adaptive Loop Filter,” IEEE International Conference on Multimedia and Expo (ICME2011), pp. 1-6, July 2011.

[6] HEVC Reference Software HM 3.0 Available from: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/branches/HM-3.0-dev/

[7] ISO/IEC JTC1/SC29/WG11, “Common test conditions and software reference configurations,” JCTVC-E700, Mar. 2011.

[8] ISO/IEC JTC1/SC29/WG11, “View synthesis algorithm in view synthesis reference software 2.0 (VSRS 2.0),” M16090, Feb. 2009.