SIMPLIFIED INTER-COMPONENT DEPTH MODELING IN 3D-HEVC

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

ABSTRACT

In this paper, we present a method to reduce complexity of depth modeling modes (DMM), which is currently used in the 3D-HEVC standardization activity. DMM adds four modes to the existing HEVC intra prediction modes; the main purpose is to accurately represent object edges in depth video. Mode 3 of DMM requires distortion calculation of all pre-defined wedgelets. The proposed method employs absolute differences of neighboring pixels in the reference block. The number of wedgelets that need to be concerned can be reduced to six. Experimental results show 3.1% complexity reduction on average while maintaining coding performance, which implies that the correct wedgelet is included, while non-viable wedgelets are disregarded.

Index Terms**—** 3D video coding, depth coding, depth modeling

1. INTRODUCTION

Compared to 2D video, 3D video adds another dimension of depth. When humans view a scene, the left eye and right eye perceive slightly different images; this is how depth perception can be achieved. Interests on 3D video have increased in the last few years due to the successes of numerous 3D video contents.

Generally, texture and depth data of multiple viewpoints are required to produce 3D video. Further, view synthesis is performed to generate texture data of an arbitrary view. Tools such as warping, interpolation and hole filling are used in this process.

In the 3D video system, the total amount of data increases in proportion to the number of views. Thus, effective data compression is vital [1, 2]. Redundancies can be observed when certain aspects are considered, e.g., temporal, inter-view and texture-depth relations. Many researchers have studied and tested such factors to develop efficient tools.

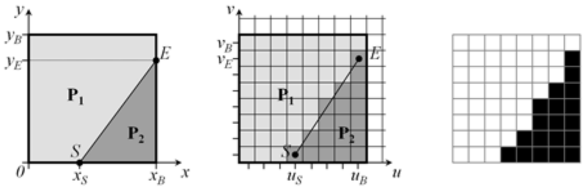
In July 2012, the moving picture experts group (MPEG) of ISO/IEC and the video coding experts group (VCEG) of ITU-T started the joint collaborative team on 3D video coding extension development (JCT3V). Numerous tools for 3D-AVC and 3D-HEVC are proposed and evaluated in the standardization activity.

2. BACKGROUND

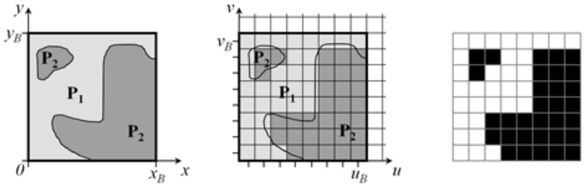
In this section, we briefly describe the depth modeling modes (DMM) currently used in 3D-HEVC. HEVC uses 35 intra prediction modes that represent angular directions. Since the conventional HEVC is for texture data compression, DMM was introduced to code depth maps more efficiently. In particular, three wedgelet modes and one contour mode are developed to represent edges in depth maps effectively [3, 4].

A wedgelet is similar to a straight line while a contour is an arbitrary-shaped connection. Wedgelet and contour partitions divides the depth block into two regions. A constant partition value (CPV) is signaled to represent its region. Figure 1 [3] exhibits an example of wedgelet and contour partitions. Figure 2 [3] displays inter-component modes, mode 3 for wedgelet and mode 4 for contour prediction. They use CTLB as the reference block.

In this paper, we focus on the mode 3 of DMM, which is the inter-component wedgelet prediction. The wedgelet for depth block is predicted from its collocated texture luma block (CTLB). Unlike 3D-AVC, 3D-HEVC does not use depth sampling. Hence, texture and depth data possess the same resolution. In other words, no interpolation is needed at CTLB in the DMM 3 process.



(a) Wedgelet partition



(b) Contour partition

Figure 1. Wedgelet and contour partitions

Table 1. Number of wedgelet candidates per block size

Depth block size	Wedgelet resolution	Number of candidates
4x4	Half-pel	86
8x8	Half-pel	782
16x16	Full-pel	1394
32x32	Double-pel	1503

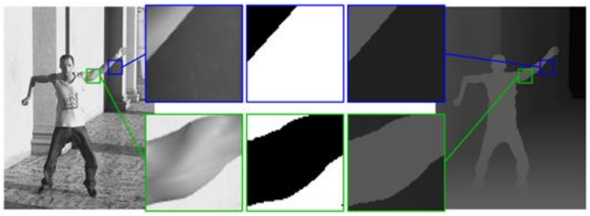


Figure 2. DMM 3 (wedgelet) and DMM 4 (contour)

Table 1 shows the number of wedgelet candidates for each depth block size. Distortions regarding each wedgelet are estimated. In the final process, the wedgelet that yields the least distortion is selected. Since the number of candidates is large, many can be removed without affecting the overall performance.

3. PROPOSED ALGORITHM

The objective of the proposed method is to simplify DMM 3 by reducing the number of required calculations. First, we apply a 1D filter to each side of CTLB to acquire four points. This filter calculates absolute differences of neighboring pixels. Instead of considering all wedgelet candidates in DMM3, we check only six candidates formed from the four acquired points [5]. Figure 3 represents the flowchart of the proposed method.

3.1. Identifying start/end points

A wedgelet is defined by a start point and an end point. Since a wedgelet is used to partition regions, the start and end points should be at the boundary of the two regions.

We calculate absolute differences of neighbors at each side of CTLB. If pixel coordinates are noted as in Figure 4, Eq. (1) is used for top and bottom sides while Eq. (2) is for left and right sides. $C_{hor,r,c}$ represents the absolute difference of its horizontally neighboring pixels for top and bottom sides where horizontal neighbors are considered. Similarly for left and right sides, $C_{ver,r,c}$ represents the absolute difference of vertically neighboring pixels.

High absolute difference indicates the intensity change at the position is large. Thus, this point is regarded as either a start point or an end point.

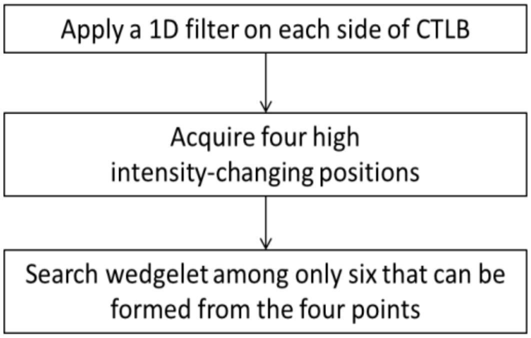


Figure 3. Flowchart of the proposed method

$A_{0,0}$	$A_{0,1}$	$A_{0,2}$	$A_{0,3}$	...
$A_{1,0}$	$A_{1,1}$	$A_{1,2}$	$A_{1,3}$	...
$A_{2,0}$	$A_{2,1}$	$A_{2,2}$	$A_{2,3}$	...
$A_{3,2}$	$A_{3,2}$	$A_{3,2}$	$A_{3,3}$	...
...	...	...	...	...

Figure 4. Pixel coordinates in CTLB

$$C_{hor,r,c} = \left|A_{r,c+1} - A_{r,c-1}\right| \tag{1}$$

$$C_{ver,r,c} = \left|A_{r+1,c} - A_{r-1,c}\right| \tag{2}$$

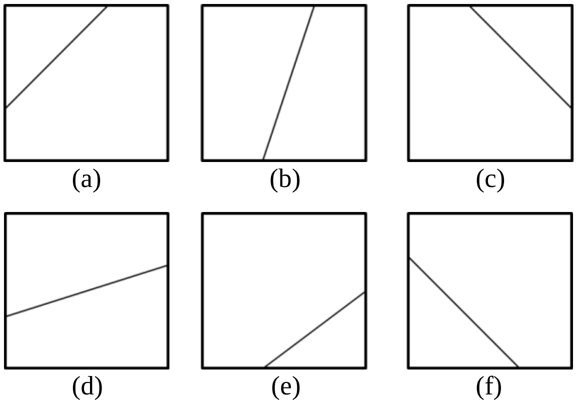


Figure 5. Wedgelet orientations (0-5)

3.2. Search in a limited set

In the previous stage, four points are acquired. With such points, six combinations wedgelets can be defined. Figure 5 exhibits the six types. In DMM wedgelet definition, the starting and end sides are pre-determined for every orientation.

Regardless of the depth block size, the proposed method reduces the number of candidates to six. The same distortion measure, sum of absolute differences (SAD), is used for wedgelet estimation. Since the number of candidates that need to be searched is significantly reduced, computational time can be saved.

To take wedgelet resolution into account, we adjust the start and end point coordinates prior to reducing wedgelet candidates. For half-pel case, i.e., 4x4 and 8x8 depth blocks, pixel positions are doubled. Similarly for double-pel case, i.e., 32x32 depth block, pixel positions are reduced by half. The coordinates are unchanged for full-pel case.

Figure 6 represents a selected wedgelet with start and end points marked by red circles. In the figure, depth intensity changes abruptly at the red circles, implying start and end points of the wedgelet. This wedgelet (orientation 0) would be selected for DMM 3 of this block; the other wedgelets in the six candidate list would generate higher distortion cost.

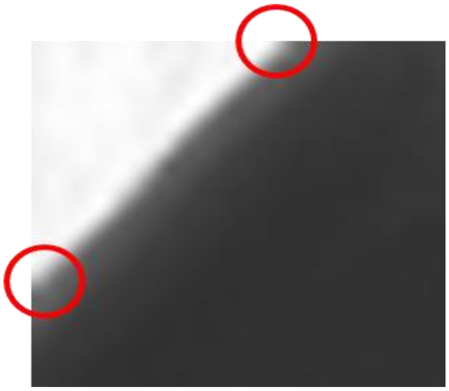


Figure 6. Identified start point and end point

Table 2. Depth coding results

Resolution	**Sequence**	**Depth QP**	**HTM-4.0.1**		**Proposed method**	
			Depth bitrate (kbps)	**Depth PSNR (dB)**	**Depth bitrate (kbps)**	**Depth PSNR (dB)**
1920×1088	Poznan_Hall2	34	64.04	39.41	64.16	39.36
		39	26.83	37.49	26.39	37.52
		42	15.30	36.68	15.39	36.67
		45	9.75	35.62	9.74	35.67
	Poznan_Street	34	116.03	41.05	116.21	41.06
		39	36.19	38.55	35.96	38.65
		42	17.88	37.02	17.93	37.11
		45	10.61	35.93	10.51	35.87
	Dancer	34	143.28	37.26	143.08	37.48
		39	67.07	35.42	67.20	36.00
		42	35.96	34.62	35.86	34.99
		45	19.12	33.41	19.28	33.38
	GT_Fly	34	166.33	32.45	166.48	32.35
		39	58.25	29.71	58.17	30.00
		42	27.85	29.27	27.62	29.24
		45	14.95	28.24	15.01	28.24
1024×768	Kendo	34	102.46	27.40	102.28	27.45
		39	37.51	24.28	37.52	24.27
		42	18.80	22.75	18.81	22.79
		45	10.78	21.90	10.82	21.89
	Balloons	34	93.26	32.53	93.51	32.51
		39	33.55	30.20	33.57	30.27
		42	17.05	28.86	17.04	28.86
		45	10.01	27.96	10.06	27.98
	Newspaper	34	145.71	33.24	145.43	33.18
		39	51.36	29.91	51.51	29.99
		42	24.34	27.39	24.40	27.56
		45	12.79	25.58	12.79	25.52
		Average BD-rate: 0.0%				

4. SIMULATION RESULTS

We have conducted experiments on “Poznan_Hall”, “Poznan_Street”, “Dancer”, “GT_Fly”, “Kendo”, “Balloons” and “Newspaper” sequences; the first four possess 1920×1088 resolutions and the other three have 1024×768 resolutions. These sequences are currently used for tool evaluation of 3D-HEVC standardization in JCT3V.

QP values of 25, 30, 35 and 40 are used for texture. For depth, 34, 39, 42 and 45 are selected from the predefined table used in the common test conditions of JCT3V [6]. Since depth maps are simpler than texture data, assigning higher QP in depth coding is commonly conducted. This is the due to the nature of depth data being much simpler than texture data.

Table 3. Decoding time comparison

Sequence	QP	Decoding time (sec)	
		HTM-4.0.1	**Proposed method**
Poznan_Hall2	25	119.42	117.78
	30	112.93	111.77
	35	110.41	109.29
	40	105.63	104.27
Poznan_Street	25	134.55	134.22
	30	127.32	127.51
	35	126.72	127.01
	40	124.04	124.59
Dancer	25	158.61	159.54
	30	151.30	150.99
	35	148.16	148.65
	40	145.45	144.23
GT_Fly	25	161.49	159.61
	30	148.98	148.19
	35	149.88	149.17
	40	145.26	144.10
Kendo	25	73.08	72.68
	30	68.31	67.77
	35	65.55	64.36
	40	61.43	60.61
Balloons	25	75.70	74.75
	30	70.97	69.56
	35	66.96	66.02
	40	63.48	62.37
Newspaper	25	64.73	64.40
	30	60.85	60.10
	35	58.55	57.53
	40	56.68	55.85
			Average reduction: 3.1%

The proposed method was implemented on HTM-4.0.1, replacing the conventional method. HTM-4.0.1 is the reference software for 3D-HEVC. Table 2 represents depth coding results of HTM-4.0.1 and the proposed method. Since the proposed method affects only depth intra coding, results for texture coding are identical, thus they are not included in the table. Bitrates and PSNR values in the table are averaged from coded views. Overall, BD-rate 0.0% implies no coding performance change. Furthermore, Table 3 shows decoding runtime comparison. The proposed method reduced the decoding runtime by 3.1% on average. Hence, the proposed method successfully simplified DMM 3 while maintaining the coding performance.

5. CONCLUSIONS

In this paper, we proposed a complexity reduction method for depth modeling mode 3 (DMM 3) currently used in 3D-HEVC. DMM modes employ wedgelets and contours to support accurate representation of sharp edges in depth blocks. For DMM 3, the corresponding texture block is used for predicting the wedgelet of depth block, yet intensive calculations are required. By calculating absolute differences, the proposed method identifies positions where high intensity changes occur in the texture block. Based on such points, the number of wedgelet candidates is reduced to six from up to 1503. Experimental results show 3.1% decoding time reduction while maintaining coding performance.

ACKNOWLEDMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012-0009228).

REFERENCES

[1] A. Smolic, "3D video and free viewpoint video—from capture to display," *Pattern recognition*, vol. 44, no. 9, pp. 1958-1968, Sep. 2011.

[2] C. Lee, B.H. Choi, and Y.S. Ho, “Efficient multiview depth video coding using depth synthesis prediction,” *Optical Engineering*, vol. 50, no. 7, July 2011.

[3] G. Tech, K. Wegner, Y. Chen, and S. Yea, "3D-HEVC test model 1," *ITU-T/ISO/IEC JCT3V*, A1005, July 2012.

[4] H. Schwarz, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, D. Marpe, P. Merkle, K. Muller, H. Rhee, G. Tech, M. Winken, and T. Wiegand, “3D video coding using advanced prediction, depth modeling, and encoder control methods,” *Picture Coding Symposium (PCS)*, pp. 1-4., May 2012.

[5] Y. Song and Y.S. Ho, "3D-CE6 related: Complexity reduction of DMM mode 3," *ITU-T/ISO/IEC JCT3V*, B0122, Oct. 2012.

[6] D. Rusanovskyy, K. Muller, and A. Vetro, "Common Test Conditions of 3DV Core Experiments," *ITU-T/ISO/IEC JCT3V*, A1100, July 2012.