SIViP (2014) 8:1031–1037
DOI 10.1007/s11760-013-0609-0

ORIGINAL PAPER

Unified depth intra coding for 3D video extension of HEVC

Yunseok Song · Yo-Sung Ho

Received: 1 September 2013 / Revised: 23 December 2013 / Accepted: 27 December 2013 / Published online: 11 January 2014

Abstract With the development of high-efficiency video coding (HEVC), the newest video coding standard, 3D video extension of HEVC (3D-HEVC) has been actively investigated. Since 3D-HEVC uses multi-view texture and depth data for input, various coding tools have been added to HEVC. In 3D-HEVC, on top of the existing 35 HEVC intra modes, eight additional modes exist, which are specifically for depth coding. In this paper, we propose a unified depth intra coding method that incorporates such depth intra modes into the regular intra mode set. In particular, the most rarely used HEVC intra modes are replaced by depth intra modes. As a result, binarization for depth intra modes is removed. Furthermore, the most probable mode selection procedure is modified to consider the elimination of several angular intra modes. The proposed method is implemented and tested on 3D video HEVC test model version 7.0. Simulation results report 2.2 % synthesis gain under all-intra configuration.

Keywords 3D-HEVC · Depth coding · Depth intra modes · Most probable mode

1 Introduction

High-efficiency video coding (HEVC) is the latest video coding standard established by the joint collaborative team on video coding (JCT-VC). Video experts from Moving Picture Experts Group (MPEG) of ISO/IEC and Video Coding Experts Group (VCEG) of ITU-T participate in this standardization [1]. HEVC is designed for superior compression of high-resolution video, targeting multimedia applications such as ultra-high definition TV (UHDTV) and mobile video services [2]. The first version of HEVC was finalized in January 2013 [3].

In March 2011, at the 96th MPEG meeting, the 3D video coding (3DVC) group of MPEG issued a call for proposals (CfP) on 3D video coding technology in two categories, advanced video coding (AVC) based and HEVC based [4]. Numerous proposals were evaluated in the 98th meeting. The most effective design at the time has been determined as a test model for development. In the following meetings, many techniques have been assessed and adopted, but the main framework has always been kept consistent. In July 2012, MPEG and VCEG held the first meeting of joint collaborative team on 3D video coding extension (JCT-3V) [5]. Their work continued the early activities of MPEG 3DVC.

3D-HEVC aims compression of multi-view texture and depth videos. While 2D video can exploit only spatial and temporal redundancies, additional aspects are considered in 3D video coding such as inter-view and texture-depth similarities [6]. Numerous 3D-HEVC-specific coding tools have been developed by taking such redundancies into account.

The basic coding structure of 3D-HEVC is depicted in Fig. 1. In a three-view coding case, temporal-wise, always the base view is coded before two dependent views; the order of dependent views depends on the video sequence. For each view, texture data are coded first and then followed by depth coding. Thus, in regard to texture coding, any modification to depth coding only affects the texture coding of dependent views, while the coded base view remains the same. In this paper, we propose a depth intra coding technique that enhances the overall coding efficiency of 3D-HEVC.

The remainder of this paper is organized as follows. In Sect. 2, depth intra coding of 3D-HEVC is described. In Sect. 3, the proposed approach is explained in detail. We analyze the simulation results in Sect. 4 and conclude the paper in Sect. 5.

Y. Song (✉) · Y.-S. Ho
Gwangju Institute of Science and Technology (GIST),
123 Cheomdangwagi-ro, Buk-gu, Kwangju 500-712, Korea
e-mail: ysong@gist.ac.kr

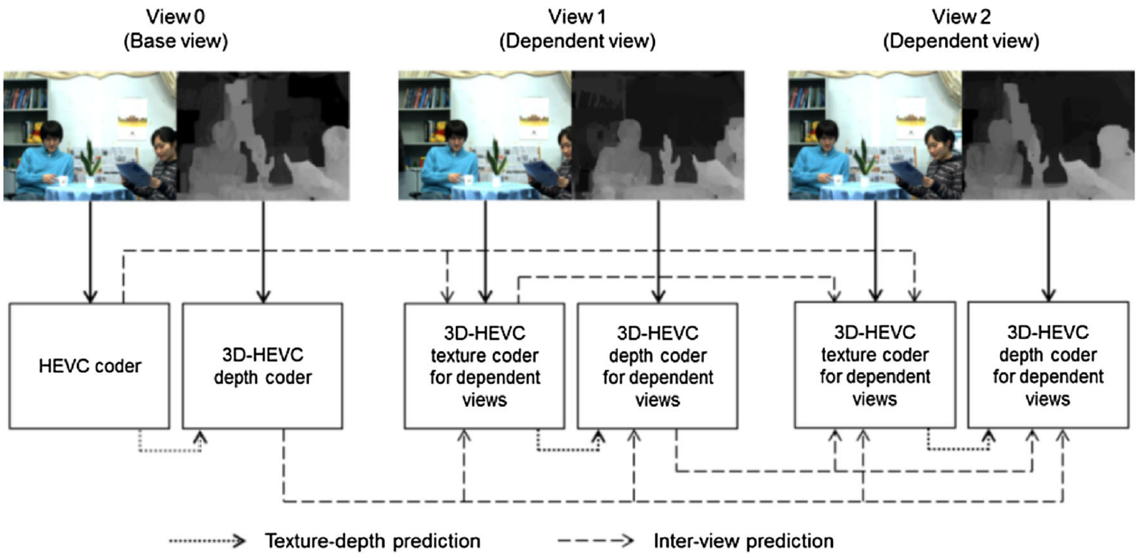


Fig. 1 Basic coding structure of 3D-HEVC

2 Depth intra coding of 3D-HEVC

In HEVC, 35 modes are used for intra prediction [7]. Figure 2 exhibits the graphical representation of such modes. Mode 0 and Mode 1 are planar and DC modes, respectively. Modes 2–34 represent 33 angular directions.

For intra coding of depth data, eight additional modes are introduced; four depth modeling modes (DMM) and one chain coding mode (CCM) are designed to accurately represent object edges in depth blocks; three simplified depth coding (SDC) modes enable residual coding [8–10]. DMM was part of various tools employed in the initial test model of 3D-HEVC; many proposals have improved its efficiency since then. CCM and SDC also have been actively investigated, and consequently, they were adopted in the first and the second JCT3V meeting, respectively. Our work is based on such methods.

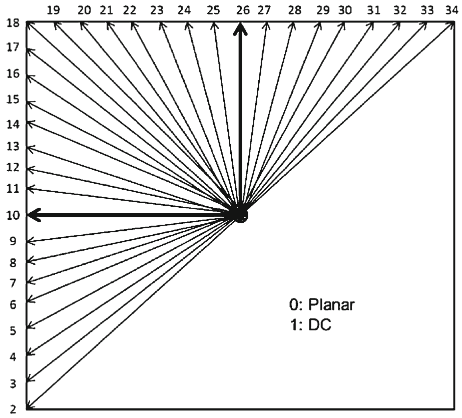


Fig. 2 HEVC intra modes

Table 1 Depth block partitioning method by each DMM

DMM 1	Explicit wedgelet signaling
DMM 2	Wedgelet partitioning from intra prediction
DMM 3	Wedgelet partitioning from colocated texture block
DMM 4	Contour partitioning from colocated texture block

Four types of DMM exist: three wedgelet prediction modes and one contour prediction mode. A wedgelet is defined as a straight line that partitions regions within a block. A contour is an arbitrary shape, which is useful when modeling curved objects. Depending on how the regions are partitioned, either a wedgelet or a contour is selected for approximation of the block. Table 1 lists how the DMMs derive partitions.

DMM 1 explicitly finds the most suitable wedgelet partition and transmits this information. In DMM 2, the wedgelet partition is predicted using neighboring blocks, which are already coded. Assuming that the line crosses several depth blocks, start and end positions of the wedgelet are relevant to those of neighboring blocks. The colocated block in texture video is also used for wedgelet and contour prediction in DMM 3 and DMM 4, respectively. As shown in Fig. 1, for each view, texture data are coded before depth data. Hence, the strong correlation between texture and depth data can be exploited.

CCM is designed to model edges of the depth block using a chain code. First, internal edges are identified based on vertical and horizontal differences between adjacent pixels. Specifically, unconnected edges are pruned and unlinked edges are connected. To code the derived edge, seven traverse types are defined: 0°, 45°, −45°, 90°, −90°, 135° and −135°. In the bitstream, the starting position of the edge and traverse codes are transmitted.

SDC is a residual coding method consisting of three modes: planar, DC and wedgelet. Planar and DC modes are based on the same modes used in the conventional intra mode set, and one segment is necessary to represent the prediction. The wedgelet mode is based on DMM 1, i.e., explicit wedgelet signaling mode. This mode is useful for representing blocks containing regions divided by a partition line. In this case, two segments are required since a wedgelet pattern is defined by a start position and an end position. Since depth data are relatively simple, such modes are sufficient to generate effective prediction. From the optimal prediction from SDC modes, for each segment, the difference between the original data and the predicted data is coded. This is also known as residual coding.

Another key feature of SDC is the use of a depth lookup table (DLT). Considering an 8-bit image, depth maps typically do not use the full range, i.e., 256 types of intensities. Thus, differences in depth data can be mapped to values using a table. In terms of bit saving, signaling the index is more beneficial than sending the residual value. A number of frames are analyzed to create the DLT, and then, this is coded in the sequence parameter set (SPS). At the decoder side, a value is reconstructed by the received residual index and DLT.

HEVC determines block sizes by comparing the RD costs from all possible block partitions. In this way, large size is selected for homogeneous regions, while small sizes are used to encode complex regions, i.e., areas where sudden change in pixel intensities occurs. Since 64 × 64 block is the largest, its region is homogeneous where object boundary is unlikely to exist. Thus, depth intra modes designed for efficient edge representation are disabled in this case; DMMs, CCM and the wedgelet mode of SDC are such modes.

The type of depth intra mode is signaled as a syntax element to differentiate them from the conventional HEVC intra modes. Table 2 represents the binarization for depth intra modes. As explained above, fewer modes are used for 64 × 64 blocks.

3 Proposed method

The objective of the proposed method is to reduce depth video bitrates while maintaining the quality. We incorporate depth intra modes into the normal HEVC intra modes. This allows binarization for depth intra modes to be unnecessary. Considering the statistical distribution of intra mode usage, rarely selected HEVC intra modes are substituted with depth intra modes. Further, most probable mode (MPM) selection procedure is modified since several angular intra prediction modes are removed.

Table 2 Binarization for depth intra modes

Depth intra mode	Bin string	
	64 × 64 CU	32 × 32/16 × 16/ 8 × 8 CU
SDC_PLANAR	0	00
HEVC intra	10	010
SDC_WEDGELET	–	011
DMM 1	–	100
DMM 4	–	101
DMM 3	–	110
SDC_DC	11	1110
DMM 2	–	11111
CCM	–	11110

3.1 Analysis of intra mode usage in depth coding

First, we analyze which HEVC intra modes are used infrequently and therefore replaceable. As shown in Table 2, eight depth intra modes exist; hence, we identify the eight least effective modes.

We conducted 3D-HEVC simulations on seven test sequences: “Poznan_Hall2,” “Poznan_Street,” “Dancer,” “GT_Fly,” “Kendo,” “Balloons” and “Newspaper” [11]. The 3D video HEVC test model (3DV-HTM) version 7.0 was used, which is the reference software for 3D-HEVC [12]. Figure 3 displays the distribution of normal HEVC intra modes in depth coding; the values are averaged from the results of seven test sequences.

The most used modes are Mode 26 (vertical) and Mode 0 (planar), occupying 24.79 and 23.57 %, respectively. On the other hand, the eight least selected modes are Modes 3, 4, 5, 6, 15, 16, 17 and 18; among these eight, Mode 15 and Mode 16 are the two most infrequent modes, chosen 0.36 and 0.30 %, respectively. As shown in Fig. 2, such modes represent directions between horizontal and diagonal.

3.2 Unified depth intra coding

In order to remove the binarization, we incorporate depth intra modes into the HEVC intra modes. In the previous section, we showed that only SDC_PLANAR and SDC_DC are used as depth intra modes when coding 64 × 64 blocks. If the block size is smaller than 64 × 64, all eight modes are used. Therefore, two sets of binarization exist.

We define two intra mode sets. For 64 × 64 block coding, Mode 15 and Mode 16 are replaced by SDC_PLANAR and SDC_DC. If the block is smaller than 64 × 64, Modes 3–6 are substituted with DMMs in the numerical order. In addition, Modes 15–18 are switched to CCM, SDC_PLANAR, SDC_WEDGELET and SDC_DC, respectively. The total

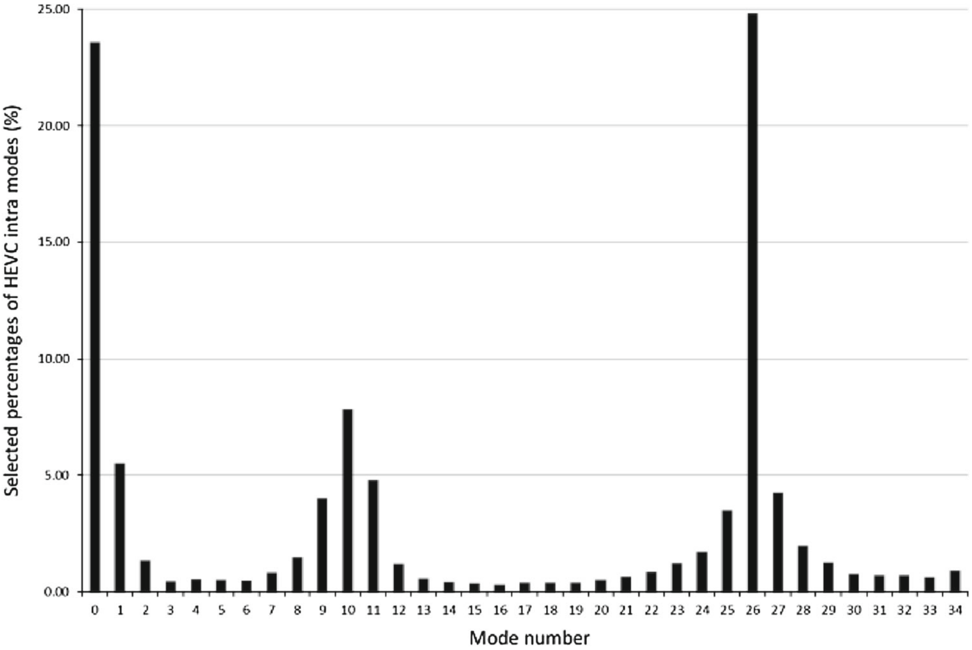
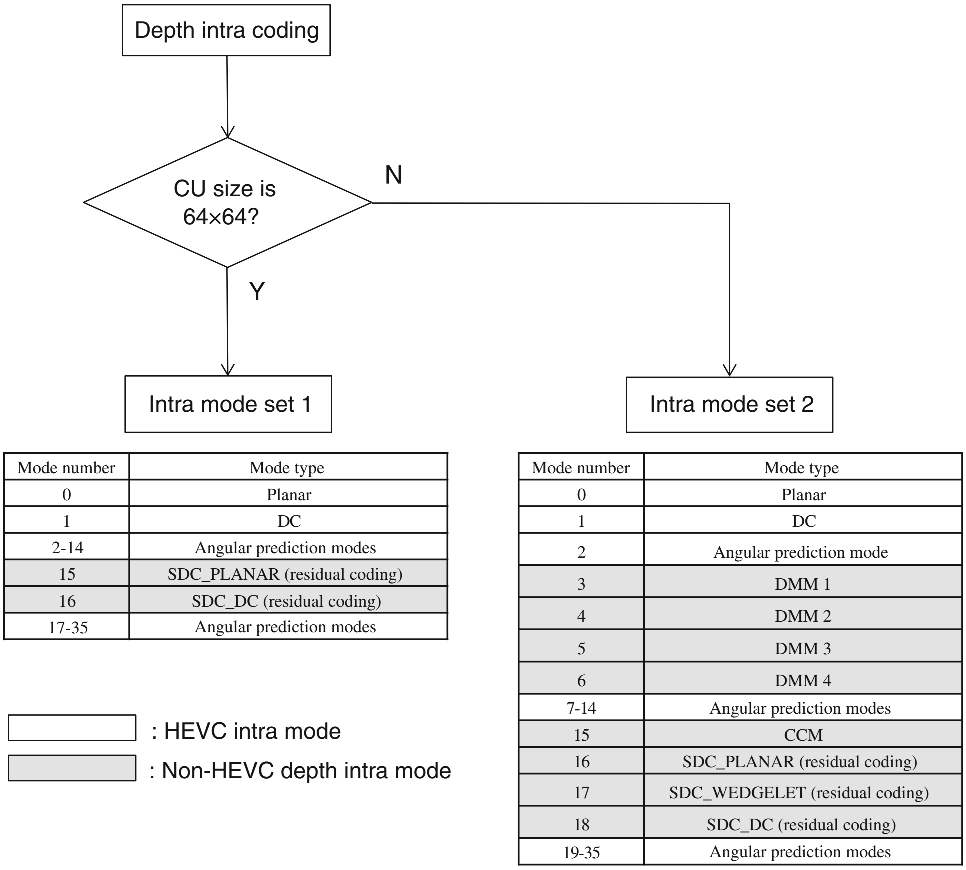


Fig. 3 Selected percentages of HEVC intra modes in depth coding



Mode number	Mode type	Shading
0	Planar	☐
1	DC	☐
2-14	Angular prediction modes	☐
15	SDC_PLANAR (residual coding)	▨
16	SDC_DC (residual coding)	▨
17-35	Angular prediction modes	☐

Mode number	Mode type	Shading
0	Planar	☐
1	DC	☐
2	Angular prediction mode	☐
3	DMM 1	▨
4	DMM 2	▨
5	DMM 3	▨
6	DMM 4	▨
7-14	Angular prediction modes	☐
15	CCM	▨
16	SDC_PLANAR (residual coding)	▨
17	SDC_WEDGELET (residual coding)	▨
18	SDC_DC (residual coding)	▨
19-35	Angular prediction modes	☐

Fig. 4 Flowchart of the proposed method

number of intra modes is 35 for both sets. The flowchart of the proposed method is displayed in Fig. 4. The shaded rows in the intra mode sets represent depth intra modes that replaced the infrequently used HEVC intra modes. If an SDC mode is to be coded, residual coding follows the coding of the mode number.

At the decoder side, first, the size of the block is checked to select the intra mode set for either 64×64 blocks or smaller

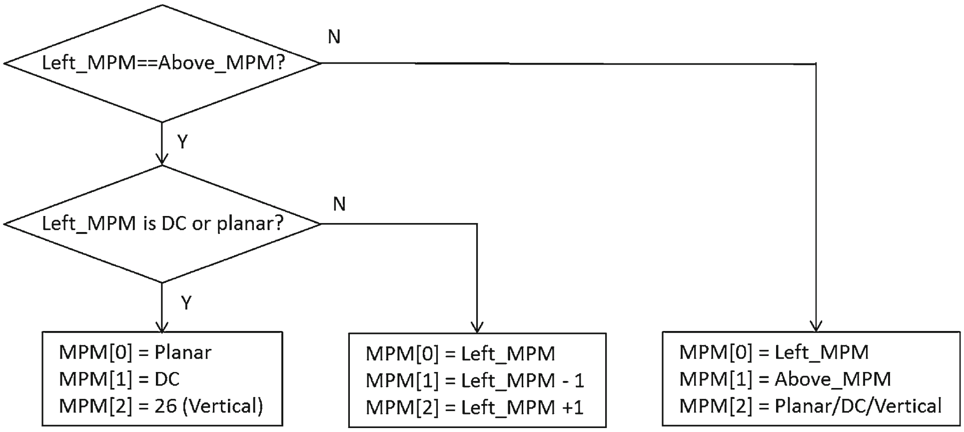


Fig. 5 MPM construction in HEVC

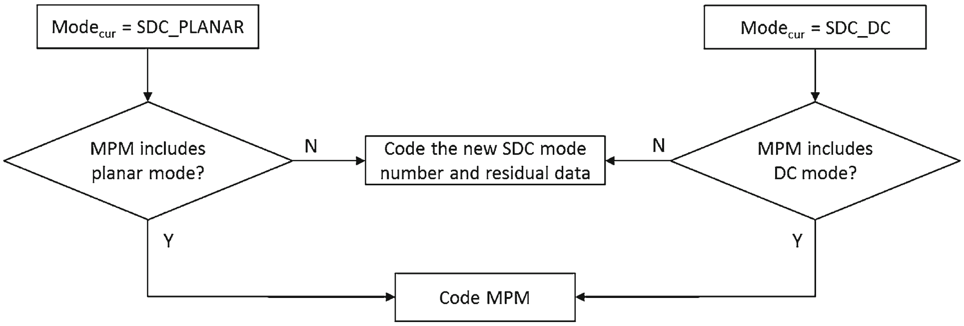


Fig. 6 MPM coding for SDC_PLANAR or SDC_DC

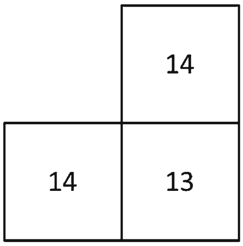
blocks. If an SDC mode is detected, decoding of residual data is executed.

3.3 MPM modification

MPMs are used to reduce bits by referencing the left and above neighboring blocks, i.e., already coded blocks. Unlike AVC, which uses a single MPM, 3D-HEVC employs three MPMs [13]. This is due to the increased number of angular directions in intra prediction.

In the proposed unified depth intra coding, SDC_PLANAR and SDC_DC are enabled for all block sizes. Rather than coding the mode number followed by residual coding, coding MPM can save many bits. As shown in Fig. 5, the probability of MPM containing HEVC planar or DC mode is very high. If the mode of the current block is SDC_PLANAR, the MPM candidate list is examined to check whether HEVC planar mode is included. Similarly for SDC_DC, we check whether HEVC DC mode exists in the MPM candidate list. If the corresponding HEVC mode is identified, its MPM is coded. In other words, the block is no longer coded as an SDC mode. This procedure is described in Fig. 6.

Since the intra mode set is differed by block size, MPM should be changed accordingly. The similarity of angular directions should be taken into account; thus, depth intra modes should not be used as an MPM candidate. If a depth intra mode is determined as an MPM candidate, the closest angular HEVC intra mode replaces this mode. This is to ensure that none of the depth intra modes are included in the MPM candidate list. Figure 7 shows an example of MPM candidate list refinement when coding a 64 × 64 block. As the original MPM candidate list includes Mode 15, which is not an HEVC intra mode, this is replaced by the closest available mode, Mode 17.



Before refinement	After refinement
MPM[0] = 13	MPM[0] = 13
MPM[1] = 14	MPM[1] = 14
MPM[2] = 15	MPM[2] = 17

Fig. 7 Example of MPM candidate refinement (64 × 64 CU)

4 Simulation conditions and results

4.1 Simulation conditions

We implemented the proposed method on 3DV-HTM 7.0. Simulations were performed on seven test sequences: “Poznan_Hall2,” “Poznan_Street,” “Dancer,” “GT_Fly,” “Kendo,” “Balloons” and “Newspaper.” The resolutions are 1920 × 1088 for the first four and 1024 × 768 for the latter three.

Table 3 Test sequences and processed views

Sequence	Resolution	Number of frames	Input views in coding order	Views to be synthesized
Poznan_Hall2	1920 × 1088	200	6-7-5	5.25, 5.5, 5.75, 6.25, 6.5, 6.75
Poznan_Street		200	4-5-3	3.25, 3.5, 3.75, 4.25, 4.5, 4.75
Dancer		250	5-1-9	2, 3, 4, 6, 7, 8
GT_Fly		250	5-9-1	2, 3, 4, 6, 7, 8
Kendo	1024 × 768	300	3-1-5	1.5, 2, 2.5, 3.5, 4, 4.5
Balloons		300	3-1-5	1.5, 2, 2.5, 3.5, 4, 4.5
Newspaper		300	4-2-6	2.5, 3, 3.5, 4.5, 5, 5.5

They possess texture data captured by a multi-view camera system and depth data estimated by means of the texture data; the exceptions are “Dancer” and “GT_Fly,” and such sequences are generated by computer graphics.

The same three views of texture and depth data are encoded. The coding order of views is center–left–right. Following the generation of output views, these are used to synthesize six intermediate views. These synthesized views are then compared with reference view synthesis data which are produced from original texture and depth data. We use the view synthesis software included in 3DV-HTM 7.0. For objective evaluation, Bjontegaard delta rate (BD-rate) is employed [14]. Bjontegaard deltas are calculated by means of average peak signal-to-noise ratio (PSNR) values and the overall bit rate. Table 3 describes the test sequences and the processed views.

We conducted experiments on all-intra and random access configurations [15]. In all-intra configuration, temporal, inter-view and texture-depth predictions are disabled. Only intra coding tools can have effect in this condition. The period of I-frame and group of picture (GOP) size are 24 and 8, respectively, for random access configuration; these values are both 1 in all-intra configuration. Naturally, videos coded under all-intra configuration result in higher qualities at increased bitrates compared to random access configuration results.

The quantization parameter (QP) values are 25, 30, 35 and 40 for texture coding. Meanwhile, higher QP values are used for depth coding: 34, 39, 42 and 45. This is due to the simple nature of depth data compared to texture data. The selection of QP values for depth coding is determined by a fixed relation table noted in [11]. Table 4 summarizes the encoder configuration.

4.2 Simulation results

The results of the proposed method are compared with the anchor results of 3DV-HTM 7.0. The performance is evaluated in terms of synthesis and texture video BD-rates based on the total bitrates. Depth video quality directly contributes to synthesized video quality that what really matters in 3D video. Thus, the depth video quality itself is not measured. Texture video quality represents the average quality of three coded texture views. Tables 5 and 6 exhibit the performances of the proposed method under all-intra and random access configurations, respectively.

Under all-intra configuration, synthesis and texture video gains are 2.2 and 1.2 %, respectively. Positive results are shown from all test sequences, 3.1 % synthesis gain for “Newspaper” being the highest. Due to the intra mode change, in terms of rate-distortion (RD) optimization, mode selection is altered. Synthesis gains show the improvement in depth video results. Since only intra coding tools are enabled, coding of texture video and depth video becomes independent. The proposed method only modifies depth intra coding. Thus, PSNR values and bitrates from texture video coding are exactly identical to those of anchor results. In other words,

Table 4 Encoder configuration

	All-intra	Random access
Period of I-frame	1	24
GOP	1	8
QP values for texture coding	25, 30, 35, 40	
QP values for depth coding	34, 39, 42, 45	

Table 5 Performance of the proposed method under all-intra configuration

Sequence	Synthesis BD-rate (%)	Texture video BD-rate (%)
Poznan_Hall2	−2.4	−1.4
Poznan_Street	−1.9	−0.8
Dancer	−1.0	−0.6
GT_Fly	−1.8	−1.1
Kendo	−2.9	−1.7
Balloons	−2.5	−1.3
Newspaper	−3.1	−1.4
Average	−2.2	−1.2

Table 6 Performance of the proposed method under random access configuration

Sequence	Synthesis BD-rate (%)	Texture video BD-rate (%)
Poznan_Hall2	−0.8	−0.6
Poznan_Street	−1.0	−0.5
Dancer	−0.4	−0.4
GT_Fly	−0.6	−0.5
Kendo	−1.0	−0.7
Balloons	−0.9	−0.6
Newspaper	−2.3	−1.1
Average	−1.0	−0.6

the 1.2 % texture video gain is the result of depth bitrate reduction.

In random access configuration, depth coding affects the quality of texture coding due to the dependencies exploited. The gains are 1.0 and 0.6 % for synthesis and texture video, respectively. They are less than the results from all-intra configuration since predictions other than intra prediction are also involved. While there is negligible change in texture bitrates, depth bitrates are vastly reduced, leading to decreased total bitrates.

5 Conclusion

In this paper, we proposed a unified depth intra coding method for 3D-HEVC. In the current 3D-HEVC, eight depth intra modes exist in addition to the 35 HEVC intra modes. Based on statistical distribution of intra mode usage, the proposed method replaces rarely selected modes with depth intra modes. Two intra mode sets are defined, depending on the block size. Due to the mode incorporation, binarization for depth intra modes is no longer necessary. Moreover, MPM is modified such that the depth intra mode cannot be an MPM candidate. The proposed method was tested on 3DV-HTM 7.0. Simulation results indicated 2.2 % synthesis gain under all-intra configuration. Hence, the proposed method successfully improved depth coding, which led to enhanced overall coding efficiency.

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. Sullivan, G.J., Ohm, J.R., Han, W.J.: Overview of the high efficiency video coding (HEVC) standard. IEEE Trans. Circuits Syst. Video Tech. **22**(12), 1649–1668 (2012)
2. Choi, J.A., Ho, Y.S.: Improved entropy coding for quantized transform coefficients in HEVC screen content coding. Signal Image Video Process. (2013). doi:10.1007/s11760-013-0542-2
3. Bross, B., Han, W.J., Ohm, J.R., Sullivan, G.J., Wang, Y.K., Wiegand, T.: High efficiency video coding (HEVC) text specification draft 10. JCTVC-L1003 (2013)
4. Video and Requirement group, “Call for Proposals on 3D Video Coding Technology”, ISO/IEC MPEG N12036 (2011)
5. Sullivan, G., Ohm, J.R.: Meeting report of the first meeting of the joint collaborative team on 3D video coding extension, development (JCT-3V). JCT3V-A1000 (2012)
6. Muller, K., Schwarz, H., Marpe, D., Bartnik, C., Bosse, S., Brust, H., Hinz, T., Lakshman, H., Merkle, P., Rhee, F.H., Tech, G., Winken, M., Wiegand, T.: 3D high-efficiency video coding for multi-view video and depth data. IEEE Trans. Image Process. **22**(09), 3366–3378 (2013)
7. Lainema, J., Bossen, F., Han, W.J., Min, J., Ugur, K.: Intra coding of the HEVC standard. IEEE Trans. Circuits Syst. Video Tech. **22**(12), 1792–1801 (2012)
8. Tech, G., Wegner, K., Chen, Y., Yea, S.: 3D-HEVC test model 4. JCT3V-D1005 (2013)
9. Heo, J., Son, E., Yea, S.: CE6.h region boundary chain coding for depth-map. JCT3V-A0070 (2012)
10. Jager, F.: 3D-CE6.h: simplified depth coding with an optional depth lookup table. JCT3V-B0036 (2012)
11. Rusanovskyy, D., Muller, K., Vetro, A.: Common test conditions of 3DV core experiments. JCT3V-D1100 (2013)
12. 3DV-HTM 7.0, https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-7.0r1
13. Wiegand, T., Sullivan, G., Bjontegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. IEEE Trans. Circuits Syst. Video Tech. **13**(7), 560–576 (2003)
14. Bjontegaard G.: Calculation of average PSNR differences between RD-curves. ITU-T VCEG-M33 (2001)
15. Jager, F.: Description of core experiment 6 (CE6) on simplified depth coding. JCT3V-D1106 (2013)