

Article

A Complexity Reduction Scheme for Depth Coding in 3D-HEVC

Qiuwen Zhang ^{1,*}, Yihan Wang ¹, Tao Wei ², Bin Jiang ¹ and Yong Gan ¹

¹ College of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China; wangyhanhenan@126.com (Y.W.); jiangbin@zzuli.edu.cn (B.J.); ganyong@zzuli.edu.cn (Y.G.)

² China National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450002, China; taoweihn@126.com

* Correspondence: zhangqwen@126.com; Tel.: +86-371-86609559

Received: 9 March 2019; Accepted: 9 April 2019; Published: 3 May 2019



Abstract: 3D-high efficiency video coding (3D-HEVC) is the next-generation compression standard for multiview system applications, which has recently been approved by MPEG and VCEG as an extension of HEVC. To improve the compression efficiency of depth map, several compression tools have been developed for a better representation depth edges. These supplementary coding tools together with existing prediction modes can achieve high compression efficiency, but require a very high complexity that restricts the encoders from ongoing application. In this paper, we introduce a fast scheme to reduce complexity of depth coding in inter and intramode prediction procedure. A simulation analysis is performed to study intra and intermode distribution correlations in the depth compression information. Based on that correlation, we exploit two complexity reduction strategies, including early SKIP and adaptive intra prediction selection. Experimental results demonstrate that our scheme can achieve a complexity reduction up to 63.0%, without any noticeable loss of compression efficiency.

Keywords: 3D video; complexity reduction; depth image

1. Introduction

In recent years, three-dimensional (3D) video has triggered an extremely attention from research institutes and industries due to the development of multimedia and stereoscopic displaying technologies [1]. A new 3D representation, multiview video plus depth (MVD), is developed by MPEG as the most popular format for 3D application [2–4]. In MVD format, only few captured texture videos and the associated depth images are utilized for providing information of the 3D scene. After receiving texture and depth data, the arbitrary virtual images can be synthesized based on depth image-based rendering (DIBR) [5]. To efficiently save the transmission cost, the texture video and depth map of MVD data are coded and transmitted. Since depth image is not displayed in receiver side and only represents the information of a 3D scene, the quality of depth map coding needs to be measured by the rendering image quality. Therefore, depth coding method designed by the new compression standard has become an active research topic.

Based on this consideration, a joint collaborative team on 3D video (JCT-3V) has been formed; 3D-high efficiency video coding (3D-HEVC) is an extension of the new compression standard HEVC [6–8]. In 3D-HEVC, several tools, such as depth modeling modes (DMMs) and segment-wise DC coding (SDC), are designed to achieve better coding of the edges in a depth map [9]. To reach the highest compression efficiency, depth coding examines all possible modes to find the one with best rate-distortion (RD) cost, which involves extremely high complexity. Thus, it is very desirable to

develop a low complexity depth coding scheme which can reduce the encoding time of 3D-HEVC encoder with no loss of compression performance.

A number of methods [10–17] have been reported to speed up the depth coding. A fast decision effort is presented in Yoon and Ho [10] to reduce the complexity of H.264/AVC. A fast depth method is proposed in Lee et al. [11] to reduce the encoding time of multiview video coding (MVC). A fast approach depending on the correlation between depth and texture is utilized in Shen et al. [12] to speed up coding computation. A fast mode decision is employed in Yoon and Ho [13] by using an adaptive edge classification that determines depth discontinuity regions. A fast motion estimation strategy is employed in Cernigliaro et al. [14] by jointly using the relation of the texture image information. A fast method is designed in Pan et al. [15] to reduce the complexity of motion estimation (ME) by utilizing the coding correlations between texture video and depth image. A fast mode decision method is introduced in Shen and Zhang [16] to minimize depth coding complexity. A joint MVD compression method has been investigated in our previous work [17] to reduce the encoding time of depth compression. All of these methods are efficient in reducing depth map complexity with acceptable video quality for the previous coding standard. However, those fast methods are not well suited to the HEVC-based 3D video encoder, which adopts a new quadtree-structured coding unit with complexity promotion in 3D-HEVC encoders.

Currently, studies on complexity reduction of depth compression have also been proposed for 3D-HEVC, which is divided to two categories: fast inter prediction and intramode decision. Category 1: fast inter prediction algorithms proposed in literatures [18–24]. An early termination method is presented in Tohidypour et al. [18] to reduce complexity of the encoders based on correlations with dependent and base view. A fast decision scheme is proposed in Shen et al. [19] to save the depth compression time by the intercomponent, interlevel, and interview correlation. However, the property of intramode prediction in depth map coding is not exploited in this method, and the interview coding mode correlation is not suitable for depth map coding; the for this reason is that most depth maps are obtained by depth estimation algorithms, and there is very little connection between depth sequence viewpoints. A low complexity method based on the motion information is developed in Zhang et al. [20] to accelerate mode decision in depth coding. An early SKIP is investigated in Conceição et al. [21] based on an adaptive threshold mode to reduce complexity of the inter frames. An early determination scheme is designed in Chen et al. [22], which can deduct an averaged 22% complexity reduction with no loss for coding. A fast decision is developed in Lei et al. [23] to reduce candidate modes of depth coding, where interview and grayscale similarity correlation are jointly used. A fast determination algorithm [24] is utilized to reduce depth encoding time. A fast mode selection is introduced in our previous work [25] to speed up the depth compression by utilizing the edge information.

Category 2: fast intra prediction algorithms [26–34]. A fast partitioning method is employed in Zhang et al. [26] to accelerate the encoding process. A fast framework is developed in Sanchez et al. [27], which is based on edge information to accelerate the depth mode prediction. Fast schemes are incorporated in Park [28,29] for balance between time reduction and RD performance of 3D-HEVC. A fast depth work [30] is proposed to reduce complexity of the RD-cost calculation. A fast depth selection is exhibited in He et al. [31] to reduce the intra candidates of the current coding unit. An intra decision scheme is proposed in Zhang et al. [32], using the squared Euclidean distance of variances to speed up the depth compression. Early termination and fast intramode selection methods are exploited in our previous works [33,34] to reduce the runtime of depth intramode process. Average complexity reductions of 40% and 30% are achieved for intra-only mode coding, respectively. These aforementioned algorithms can effectively accelerate intra and intermode decision with acceptable video quality degradation. However, useful depth map compression information including the intermode, intramode, and RD cost correlation among 3D-HEVC encoders are not fully exploited in the above algorithms. There is still some room for further reducing computational complexity of depth coding.

In this paper, we present a novel 3D-HEVC scheme to reduce mode prediction complexity of depth coding. The basic idea of this scheme is to use the spatial–temporal and texture video correlations to analyze current treeblock characteristics and to skip some unnecessary modes. It consists of two methods: early SKIP and adaptive intra prediction selection.

The rest of the paper is organized as follows. Section 2 presents the mode statistical analysis of depth map coding. Section 3 explains the proposed fast scheme in detail. The experimental results and conclusion are provided in Sections 4 and 5, respectively.

2. Mode Statistical Analysis

To better understand the depth map characteristics, 3D-HEVC coder mode prediction, and computational complexity distribution we conducted several statistical studies.

2.1. Test Experimental Setup

The test experiment is based on 3D-HEVC (HTM 16.1 [35]) coding results over 8 sequences jointly selected by the JCT-3V. Among those sequences, “Undo_Dancer”, “GT_Fly”, “Poznan_Street”, “Poznan_Hall2”, and “Shark” are 1920 × 1088 resolution, while “Kendo”, “Balloons”, and “Newspaper” are 1024 × 768 resolution [36]. Test conditions are listed as 3-view case, quantization parameters (QPs), chosen with 20, 25, 35, and 45; number of coding tree units is 64; CU depth level is from 0 to 3; DMMs and VSO are enabled; and the number of test frames is 150. It should be noted the test experimental setup in Section 2 uses QPs values different from those used in 3D-HEVC CTC. This is because that those QPs values are better to find out the intra- and interframe mode distribution changes.

2.2. Inter Prediction Mode Distribution

Firstly, an intermode distribution between different QPs is investigated by using the HTM 16.1 in Table 1. It can be observed that most treeblocks in inter coding choose SKIP as the best mode, and the percentage of SKIP in depth map treeblocks increases as the QP increases. The percentage of choosing SKIP is 83.4%, while the percentages for a depth map treeblock to be coded with Merge, 2N × 2N, Intra, and other intermodes are very low, no more than 5.1%. The reason is that depth sequences have more stationary area and this type of region mostly chooses the SKIP mode. Especially for small global motion sequences such as “Poznan_Hall2”, the choice of SKIP mode is very high, i.e., 98.7% in QP 45. Thus, if we can predetermine for a depth map treeblock whether the optimal intermode is SKIP or not, and a lot of computation can be reduced in depth map coding.

2.3. Mode Distribution in Intra Prediction

The statistical result of intramode distribution for depth map is also shown in Table 2. All statistical studies are defined under intra-only configuration. It could be observed that most treeblocks after depth intra prediction choose planar mode, while the percentage of choosing planar as the best mode is 74.0%. Especially for “Poznan_Hall2”, coded in QP = 45, the 97.1% treeblock of depth map is coded with Planar mode. The reason is that stationary areas of depth map mostly select the Planar mode. Although “Shark” and “Undo_Dancer” contain a complex motion, approximately 32.1% to 88.2% of treeblocks of the depth map still choose Planar mode. It can be observed that the number of treeblocks choosing Planar mode increases when QP increases. Therefore, if we can decide a treeblock of depth map whether the optimal mode is the Planar or not, the wasteful intramode process can be omitted, and thus a very time-consuming part of the computation can be saved. Moreover, the probability for a treeblock of depth map to be coded with DMMs is ~2.2%. However, the DMMs introduce a drastic increase in complexity for depth map coding. Thus, it is necessary to develop a fast scheme to remove the redundant RD calculation of DMMs.

Table 1. Intermode distribution in depth compression.

Sequences	QP	SKIP Mode	Merge Mode	Inter 2N × 2N	Intramodes	Other Intermodes
Kendo	20	58.9%	3.9%	12.9%	15.8%	8.5%
	25	83.7%	2.1%	5.2%	3.2%	5.8%
	35	91.9%	1.2%	2.9%	1.6%	2.4%
	45	95.8%	0.7%	1.6%	1.1%	0.8%
Balloons	20	64.7%	5.2%	9.1%	9.3%	11.7%
	25	85.6%	3.1%	4.6%	3.9%	2.8%
	35	92.7%	1.8%	2.2%	1.9%	1.4%
	45	97.0%	1.1%	0.8%	0.6%	0.5%
Newspaper	20	56.7%	7.1%	9.2%	8.1%	18.9%
	25	80.9%	4.3%	5.4%	4.2%	5.2%
	35	90.6%	2.9%	3.1%	1.9%	1.5%
	45	94.7%	1.7%	1.2%	1.1%	1.3%
Shark	20	53.1%	11.8%	8.1%	7.3%	19.7%
	25	76.8%	8.9%	3.3%	3.2%	7.8%
	35	88.1%	5.6%	1.2%	1.4%	3.7%
	45	93.7%	2.5%	0.6%	0.7%	2.5%
Undo_Dancer	20	56.2%	10.9%	7.9%	7.1%	17.9%
	25	78.6%	7.8%	4.0%	2.6%	7.0%
	35	89.5%	5.2%	1.5%	1.1%	2.7%
	45	93.9%	2.3%	0.9%	0.6%	2.3%
GT_Fly	20	68.7%	5.2%	6.1%	11.7%	8.3%
	25	86.3%	3.3%	3.5%	4.9%	2.0%
	35	93.1%	1.2%	2.1%	2.6%	1.0%
	45	97.4%	0.8%	0.7%	0.6%	0.5%
Poznan_Street	20	67.1%	5.3%	6.7%	9.6%	11.3%
	25	84.9%	3.2%	3.6%	4.7%	3.6%
	35	92.2%	1.8%	2.1%	2.3%	1.6%
	45	97.5%	0.7%	0.6%	0.5%	0.7%
Poznan_Hall2	20	72.6%	4.1%	6.2%	9.3%	7.8%
	25	90.6%	2.2%	2.5%	2.7%	2.0%
	35	95.8%	1.1%	1.2%	0.9%	1.0%
	45	98.7%	0.2%	0.3%	0.4%	0.4%
Average		83.4%	3.7%	3.8%	4.0%	5.1%

Table 2. Intramode distribution for depth map.

Sequences	QP	Planar(0) Mode	DC(1) Mode	Horizontal(10) Mode	Vertical(26) Mode	Other HEVC Intramodes	DMMs
Kendo	20	35.8%	7.9%	15.6%	21.8%	6.5%	5.4%
	25	69.9%	2.3%	7.1%	11.3%	2.6%	1.1%
	35	78.9%	1.1%	5.7%	8.8%	1.4%	0.5%
	45	91.2%	0.6%	2.1%	3.4%	14.8%	0.2%
Balloons	20	63.9%	5.9%	8.8%	12.3%	7.9%	3.7%
	25	78.3%	2.7%	4.9%	7.6%	5.2%	1.8%
	35	87.9%	1.3%	3.4%	4.8%	2.4%	0.9%
	45	93.6%	0.8%	1.5%	2.7%	23.5%	0.4%
Newspaper	20	46.1%	6.8%	13.6%	18.7%	15.3%	4.6%
	25	73.2%	2.7%	5.8%	10.4%	3.3%	1.3%
	35	83.6%	1.3%	3.5%	6.4%	2.8%	0.8%
	45	92.2%	0.9%	1.8%	2.7%	24.6%	0.5%
Shark	20	32.1%	4.8%	16.8%	22.8%	14.8%	3.9%
	25	57.2%	3.6%	9.6%	14.3%	3.5%	2.2%
	35	78.9%	2.3%	6.4%	9.1%	2.8%	1.4%
	45	87.6%	1.2%	3.5%	4.9%	17.7%	0.8%
Undo_Dancer	20	33.2%	4.6%	15.9%	21.7%	8.4%	4.3%
	25	59.6%	3.8%	8.7%	13.1%	3.6%	1.9%
	35	79.6%	2.5%	5.9%	8.5%	4.9%	1.3%
	45	88.2%	1.5%	3.2%	4.4%	14.8%	0.9%
GT_Fly	20	43.9%	3.7%	15.8%	18.9%	5.0%	2.9%
	25	72.4%	4.8%	6.1%	8.3%	1.3%	9.4%
	35	86.5%	1.9%	3.1%	4.9%	1.1%	8.9%
	45	90.1%	1.1%	1.6%	2.3%	10.7%	1.7%
Poznan_Street	20	49.8%	7.1%	12.7%	15.6%	3.4%	3.9%
	25	81.2%	2.5%	4.8%	6.5%	2.3%	1.2%
	35	92.7%	2.4%	1.5%	2.1%	0.9%	0.5%
	45	96.2%	0.8%	0.7%	1.2%	6.5%	0.2%
Poznan_Hall2	20	65.1%	5.1%	7.8%	11.3%	2.6%	2.4%
	25	86.7%	2.1%	3.2%	4.6%	1.4%	1.3%
	35	94.7%	1.1%	0.8%	1.1%	14.8%	0.3%
	45	97.1%	0.8%	0.5%	0.7%	7.9%	0.1%
Average		74.0%	2.9%	6.3%	9.0%	7.5%	2.2%

Other HEVC intramodes are Angular modes 2–9, 11–25, and 27–34.

Similar to HEVC, a computationally expensive exhaustive mode decision is computed using all the possible inter and intramodes to find the least RD cost for each treeblock in 3D-HEVC. Furthermore, additional new types of coding tools have been added to 3D-HEVC for improving the compression efficiency of depth map. Those coding tools achieve the highest possible coding efficiency, but also lead to extreme computational complexity. Therefore, fast mode decision algorithms, which can reduce 3D-HEVC computational complexity without compromising coding efficiency, are very desirable for real-time implementation of 3D-HEVC encoders.

3. Proposed Complexity Reduction Algorithm Design

3.1. Early SKIP Mode Decision

In depth map compression, the interframe prediction of 3D-HEVC can be coded with SKIP, Merge, normal inter, and intramodes. When a treeblock is coded with SKIP mode, it has no reference picture index and coded motion vector delta. From this, once SKIP mode is predecided, variable intermode computation for current CU can be significantly reduced in depth map coding. Since depth maps have large areas of nearly constant and homogeneous regions compared to texture video, many treeblocks of the depth map end up being decided as SKIP mode after computing the RD costs of all prediction modes in 3D-HEVC encoders. As described in Section 2, if we could find a proper condition to early distinguish the treeblock of depth map from corresponding treeblocks, which is more likely to select SKIP mode. The intermode decision can be effectively simplified and large computational complexity can be skipped in 3D-HEVC encoders.

In depth sequences, there exist high correlations between the neighboring treeblocks in spatial and temporal domains [37]. The mode information of the current treeblock is very close to its spatial and temporal adjacent treeblocks. Furthermore, strong coding information correlations can be found in depth map and texture video. Therefore, we can use of the corresponding spatial–temporal and texture video correlations to analyze current treeblock characteristics and skip some unnecessary intermodes, which are rarely used in adjacent treeblocks.

According to these observations, we analyze current treeblock by utilizing the information correlations from spatial–temporal and the corresponding texture treeblocks. A set of adjacent predictors (π) is defined as follows

$$\pi = \{B_1, B_2, B_3, B_4, B_5, B_6\} \quad (1)$$

where B_1, B_2, B_3 , and B_4 denote spatial predictors in the current depth map view in Figure 1, B_5 denotes temporal predictor in the previously coded depth map frame for current treeblock T_C ; B_6 denotes the texture predictor from corresponding texture video treeblock.

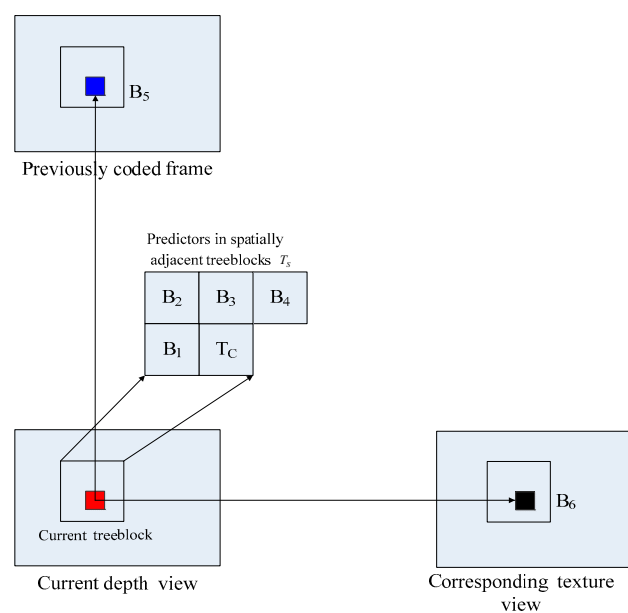


Figure 1. Predictors of the treeblock in depth view.

According to information from the predictors π , the current treeblock can be extracted to adaptively skip unnecessary intermodes. To analyze the correlation, the current treeblock can be classified into two types, “DI” and “DII” are defined as follows.

“DI” represents the spatial–temporal and the corresponding texture video treeblocks of current depth map treeblock ($B_1, B_2, B_3, B_4, B_5,$ and B_6) are that all choose on SKIP mode.

“DII” denotes remaining treeblocks.

Figure 2 gives the SKIP distribution for two depth treeblocks. It is observed that for type of “DI”, SKIP is more like selected than the other intermodes, the probability of select the SKIP is larger than 98%, and the probability of select other modes is less than 2%. For type “DII” in Figure 2b, the probability of SKIP is 34.9% in inter coding. In addition, we give the statistical analysis of the average ratio of type “DI” treeblocks in Figure 3. It can be observed that most treeblocks in inter coding belong to the type “DI”, ~75.1%. Even the sequence with complex motion as “Undo_Dancer”, more than 69.7% treeblocks belongs to “DI”. If the depth coding can decide SKIP for “DI” at the early stage, 75.1% inter encoding time will be reduced, and the computational complexity of depth compression can be effectively saved. Based on the above analysis, if the treeblock belong to type “DI”, choose the SKIP as the best mode.

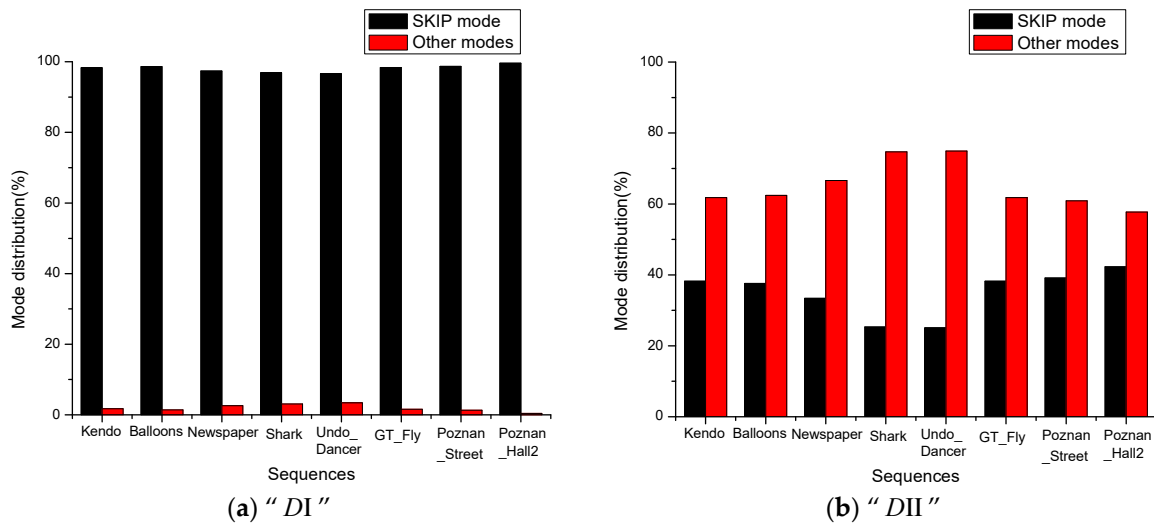


Figure 2. SKIP mode distribution for each type of treeblock.

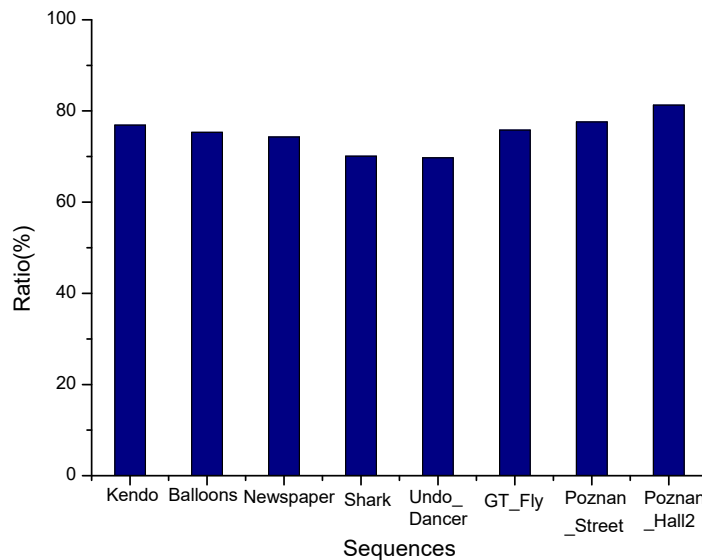


Figure 3. Statistical analysis of the ratio of “DI” treeblocks.

3.2. Adaptive Intra Prediction Selection

At each depth level, both the HEVC intramodes and DMMs are used to select the best one with least RD cost. This is a significant runtime-consuming part due to high computation of cost calculation for 3D-HEVC. In fact, both planar and DC modes are good indicators of smooth regions. On the other hand, planar and DC modes are rarely chosen for treeblocks with complex texture region. Moreover, DMMs are designed for depth images with sharp transition and are inefficient for flat region coding. Since depth image contains large areas of varying sample values, most of the DMMs predict can be omitted in the smooth region. Based on this consideration, if the optimal intramode can be decided in advance, the consuming part due to full RD cost calculation can be skipped, and thereafter a lot of computation can be reduced in depth compression. Therefore, we propose an adaptive intra prediction selection for 3D-HEVC by utilizing the complexity of depth block.

A mode complexity parameter (*IMP*) of depth treeblock is defined on the spatial–temporal and texture video treeblocks in predictors (π) as follows

$$IMP = \frac{1}{N} \cdot \sum_{i \in \pi} K_i \cdot \theta_i \cdot \rho_i \tag{2}$$

where N denotes the number of treeblocks in predictors π (including $B_1, B_2, B_3, B_4, B_5,$ and B_6) and K_i denotes the weight factor, which is the correlation of the adjacent treeblocks in predictors π . Based on correlation, the value of K_i is summarized in Table 3. θ_i denotes the mode weight factor of predictor, which is dependent on the mode complexity of neighboring treeblocks: when B_i is Planar and DC mode, θ_i is set to “1”; when the predictor B_i is Horizontal and Vertical mode, and θ_i is set to “2”; when the predictor B_i is Angular 2-9, Angular 11-25, and Angular 27-34, and θ_i is set to “3”; and when the predictor B_i is DMMs and θ_i is set to “4”. ρ_i denotes the adjust parameter. Hence, when B_i is available, ρ_i is considered to be “1”; otherwise, ρ_i is considered to be “0”.

Table 3. The K_i assigned to each predictor.

Index (i)	B_1	B_2	B_3	B_4	B_5	B_6
K_i	0.2	0.15	0.2	0.15	0.15	0.15

Generally, more complex depth map treeblocks have a larger value in Equation (2). Based on the value of *IMP*, the current depth map treeblock in 3D-HEVC can be divided into simple, normal, and complex regions.

$$\begin{cases} IMP < Thr_0 & Treeblock \in simple\ region \\ Thr_0 \leq IMP \leq Thr_1 & Treeblock \in normal\ region \\ IMP > Thr_1 & Treeblock \in complex\ region \end{cases} \tag{3}$$

where Thr_0 and Thr_1 are the mode weight factors. According to the extensive experiments, the value of Thr_0 and Thr_1 are set to 1.2 and 2.8, respectively. These values achieve a good and consistent performance on a variety of test sequences with different texture characteristics and motion activities, and fixed for each treeblock QP level in 3D-HEVC encoders.

To analyze the intramode correlation in depth map coding, we give the intramode distribution for different treeblocks in HTM encoders. It can be seen from Table 4 that for the treeblock with a simple region, ~99.5% of the depth map’s CUs select the optimal mode in Planar mode (97.0%) and DC mode (2.5%), while the average probability of remaining intramodes have a very low probability (0.5%). For the “Poznan_Hall2”, the average probabilities of choosing Other HEVC intramodes and DMMs is small (0.0%), because it contains a large static background. Therefore, it can be observed from the Table 4 that the block in simple region only needs to perform Planar and DC modes. For the block in normal region on Table 5, the probability of choosing Planar, DC, Horizontal, and Vertical mode is 69.9%, 2.3%, 10.3%, and 13.2%, respectively, and the probability of remaining intramodes is

less than 4.3%. The total probability of Planar, DC, Horizontal, and Vertical mode is ~95.7%, and thus it is not necessary to select DMMs, Angular 2-9, 11-25, and 27-34. For the treeblock with a complex region in Table 6, the probability of choosing Planar, DC, Horizontal, Vertical, Other HEVC intramodes, and DMMs is in total more than 5.2%, which is not negligible. Thus, the treeblock with a complex region needs to perform all intramodes. According to the above analysis, the optimal intramodes that will be performed in depth treeblocks are described in Table 7.

Table 4. Intramode distributions for simple region of treeblocks in depth map.

Sequences	Planar(0) Mode	DC(1) Mode	Horizontal(10) Mode	Vertical(26) Mode	Other HEVC Intramodes	DMMs
Kendo	97.2%	2.4%	0.1%	0.2%	0.1%	0.0%
Balloons	98.1%	1.5%	0.1%	0.2%	0.1%	0.0%
Newspaper	96.4%	3.1%	0.2%	0.2%	0.1%	0.0%
Shark	95.6%	3.7%	0.1%	0.1%	0.2%	0.3%
Undo_Dancer	95.4%	3.5%	0.1%	0.2%	0.4%	0.4%
GT_Fly	95.8%	3.3%	0.1%	0.2%	0.3%	0.3%
Poznan_Street	98.3%	1.5%	0.1%	0.3%	0.0%	0.0%
Poznan_Hall2	99.2%	0.7%	0.0%	0.1%	0.0%	0.0%
Average	97.0%	2.5%	0.1%	0.2%	0.1%	0.1%

Table 5. Intramode distributions for normal region of treeblocks in depth map.

Sequences	Planar(0) Mode	DC(1) Mode	Horizontal(10) MODE	Vertical(26) MODE	Other HEVC Intramodes	DMMs
Kendo	72.9%	2.4%	9.9%	12.1%	2.2%	0.5%
Balloons	74.2%	2.1%	8.8%	11.5%	3.1%	0.3%
Newspaper	70.6%	3.2%	10.4%	13.2%	2.0%	0.6%
Shark	61.7%	1.9%	13.1%	16.5%	5.7%	1.1%
Undo_Dancer	61.4%	1.7%	12.9%	16.2%	6.5%	1.3%
GT_Fly	63.2%	2.1%	12.3%	15.4%	6.1%	0.9%
Poznan_Street	75.7%	2.7%	7.9%	10.8%	2.6%	0.3%
Poznan_Hall2	79.2%	2.6%	6.8%	9.7%	1.6%	0.1%
Average	69.9%	2.3%	10.3%	13.2%	3.7%	0.6%

Table 6. Intramode distributions for complex region of treeblocks in depth map.

Sequences	Planar(0) Mode	DC(1) Mode	Horizontal(10) Mode	Vertical(26) Mode	Other HEVC Intramodes	DMMs
Kendo	49.1%	6.9%	5.3%	7.2%	13.6%	17.9%
Balloons	52.4%	5.8%	6.1%	8.3%	12.6%	14.8%
Newspaper	47.2%	7.4%	6.5%	9.7%	13.3%	15.9%
Shark	39.2%	5.5%	3.7%	5.3%	8.9%	37.4%
Undo_Dancer	38.7%	5.8%	3.5%	4.9%	8.6%	38.5%
GT_Fly	40.4%	6.4%	4.4%	5.5%	9.1%	34.2%
Poznan_Street	53.9%	5.8%	5.8%	7.9%	11.7%	14.9%
Poznan_Hall2	56.7%	7.6%	6.2%	9.2%	8.6%	11.7%
Average	47.2%	6.4%	5.2%	7.3%	10.8%	23.2%

Table 7. Optimal intramode for each type treeblocks.

Depth Map Type	Candidate Intramodes
Treeblocks in simple region	Planar and DC mode
Treeblocks in normal region	Planar, DC, Horizontal, and Vertical mode
Treeblocks in complex region	All intramodes

3.3. Overall Algorithm

On the basis of the above analysis, the proposed scheme including strategies of early SKIP and adaptive intra prediction selection are summarized as follows.

Step 1: Start mode decision with a treeblock.

Step 2: Locate the information of the spatial–temporal and the corresponding texture video treeblocks in predictors π .

Step 3: Perform early SKIP mode decision. If the predictors in π are all choose SKIP, only SKIP used for best mode, and go to Step 6.

Step 4: Test adaptive intra prediction selection. Compute IMP in Equation (5) and classify the treeblock into three regions.

Step 5: When depth treeblock is in simple region, the optimal intramodes are Planar and DC; when depth treeblock is in the normal region, the optimal intramodes are Planar, DC, Horizontal, and Vertical; otherwise, all intramodes are performed.

Step 6: Select the best mode.

4. Experimental Results

In order to evaluate the proposed complexity reduction scheme, the experiment is implemented on the 3D-HEVC (HTM 16.1 [35]). Eight MVD sequences, recommended by JCT-3V CTC [36], are used for the simulation. Among those test sequences, the “Shark”, “Undo_Dancer”, and “GT_Fly” are synthetic videos with high precision depth map, and the “Kendo”, “Balloons”, “Newspaper”, “Poznan_Hall2”, and “Poznan_Street” are natural videos with estimated depth map. The experiment conditions: three-view: center view-left view-right view (in coding order); QPs are set as (25, 34), (30, 39), (35, 42), and (40, 45); the number of test coding frames is 100; DMMs Mode is on; and VSO is enabled. The “VSRS-1D-Fast” software is used for experiments [38]. We use the Bjontegaard

Delta Bitrate (BDBR) [39] to evaluate the compression performance of our scheme. The complexity is measured with reduction of total time “ TS_D ”, which is defined as follows

$$TS_D = \frac{Time_{proposed} - Time_{Anchor}}{Time_{Anchor}} \times 100\% \quad (4)$$

where $Time_{proposed}$ and $Time_{Anchor}$ represent the run time of proposed encoder scheme and anchor method for same test sequence, respectively. The coding performances were implemented on a Microsoft Windows 7 SP1 (64-bit) workstation with two CPU of E5-2640@2.0 GHz.

Table 8 gives the experiment results of the proposed scheme compared to the HTM16.1 encoder. $BDBR_S$ and $BDBR_T$ represent the BD-rates calculated by the synthesized view and texture video PSNR, respectively. It can be observed from Table 8 that the proposed scheme can significantly save the run time for both all-intra and random access configurations. The proposed scheme saves 49.1% and 63.0% encoding time for “Random Access” and “All-Intra” configurations, respectively, with a minimum of 45.6% for “Shark” (“Random Access” case) and a maximum of 68.7% for “Poznan_Hall2” (“All-Intra” case). The encoding time saving of “Poznan_Hall2” sequence is more than that of other test videos; this is because that the “Poznan_Hall2” sequence contains more stationary areas. The computation reduction of the proposed scheme is particularly high because a lot of unnecessary modes are considerably reduced. Meanwhile, the proposed scheme slightly reduces the quality of rendered video because the depth image has small loss. The average increase of bitrate on $BDBR_S$ is less than 0.29% and 0.24% for “Random Access” and “All-Intra” configurations, respectively. Additionally, the proposed scheme can improve the bitrate performance on $BDBR_T$ because the bitrate of depth map is reduced. It leads to 0.12% (“Random Access” case) and 0.07% (“All-Intra” case) BDBR decrease compared with original HTM16.1 method. Thus, the proposed scheme can effectively save the run time of depth coding with a little loss of coding efficiency.

Table 8. Coding results of the proposed scheme.

Sequences	Random Access			All-Intra		
	$BDBR_T$ (%)	$BDBR_S$ (%)	TS_D (%)	$BDBR_T$ (%)	$BDBR_S$ (%)	TS_D (%)
Kendo	−0.14	0.36	−50.7	−0.16	0.29	−64.8
Balloons	−0.19	0.18	−49.2	0.03	0.14	−61.6
Newspaper	−0.27	0.21	−47.9	0.02	0.34	−60.9
Shark	0.02	0.32	−45.6	0.04	0.28	−59.7
Undo_Dancer	0.04	0.29	−46.4	0.03	0.26	−59.5
GT_Fly	0.03	0.11	−47.6	−0.12	0.05	−63.8
Poznan_Street	−0.17	0.26	−51.5	−0.18	0.18	−65.2
Poznan_Hall2	−0.31	0.57	−53.8	−0.23	0.39	−68.7
Average	−0.12	0.29	−49.1	−0.07	0.24	−63.0

Figures 4 and 5 show the RD performance and depth map runtime saving curves of the proposed scheme compared to original encoder on two test sequences “Kendo” (1024 × 768 resolution with natural sequence) and “Undo_Dancer” (1920 × 1088 resolution with synthetic sequence) under “Random Access” and “All-Intra” configurations. It can be observed from Figures 4 and 5 that the proposed scheme can achieve a consistent depth map encoding time saving over a large QP range with nearly the same value in the rendered view PSNR. Furthermore, when encoding bit rate decreases, the proposed scheme time saving increases, as seen in Figures 4 and 5. This is because as QPs increase, the probability of checking SKIP due to early SKIP mode decision algorithm and the probability of testing Planar/DC mode due to adaptive intra prediction selection algorithm are all increased in the proposed scheme.

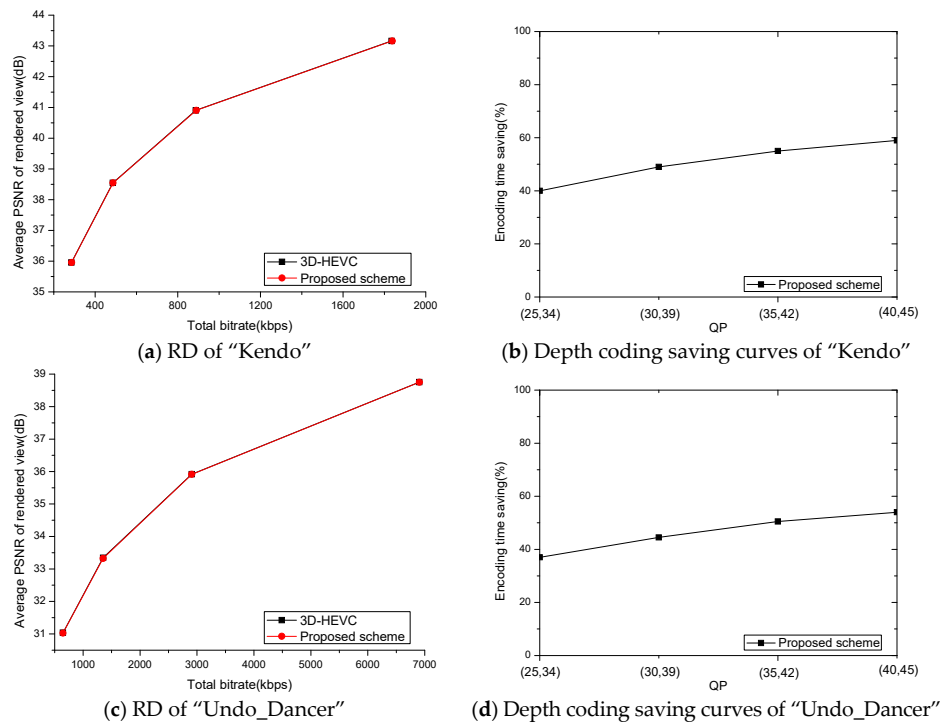


Figure 4. RD performance and depth map time saving under different QPs (25, 34), (30, 39), (35, 42), and (40, 45) for the "Random Access" condition.

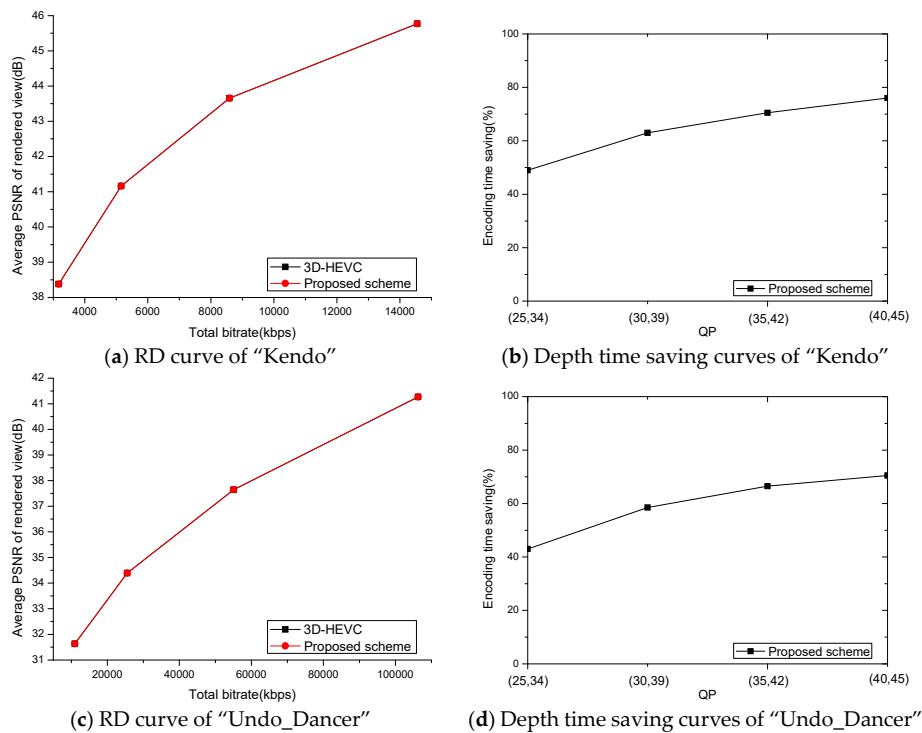


Figure 5. RD performance and depth map time saving under different QPs (25, 34), (30, 39), (35, 42), and (40, 45) for the "All-Intra" condition.

In addition to the original HTM, the result of proposed scheme is also compared with five state-of-the-art fast methods under the "Random Access" and "All-Intra" conditions. These methods include FMDRA [19] (optimized only for depth map inter coding), CRSED [27], EBIMS [29], LBPMS [30], PDIMS [32], and EETDM [33], which are recently developed fast and efficient depth map algorithms

for 3D-HEVC. All algorithms were performed on HTM16.1 and tested using the same computer for objective comparison. It can be seen from Figures 6 and 7 that all five previous works have good encoding time saving. Compared with the five previous algorithms, the proposed scheme can perform better performance on computation reduction. Approximately 8.5%–41.0% depth runtime can be further saved in the “Random Access” and “All-Intra” conditions. On the other hand, compared with CRSED and EBIMS methods, the coding efficiency loss of the proposed scheme is negligible, with less than 0.2% BDBR_s increase. Furthermore, the proposed scheme performs a better RD performance than FMDRA, LBPMS, PDIMS, and EETDM methods. The above results show that the proposed method outperforms the recent fast algorithms for 3D-HEVC, in terms of depth coding runtime saving with better or similar RD performance.

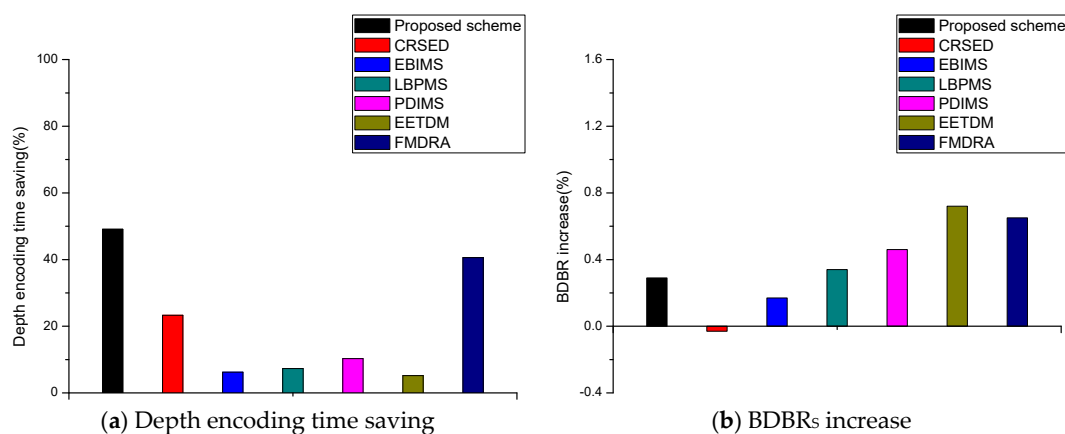


Figure 6. Coding results of the proposed scheme and five algorithms under the “Random Access” configuration.

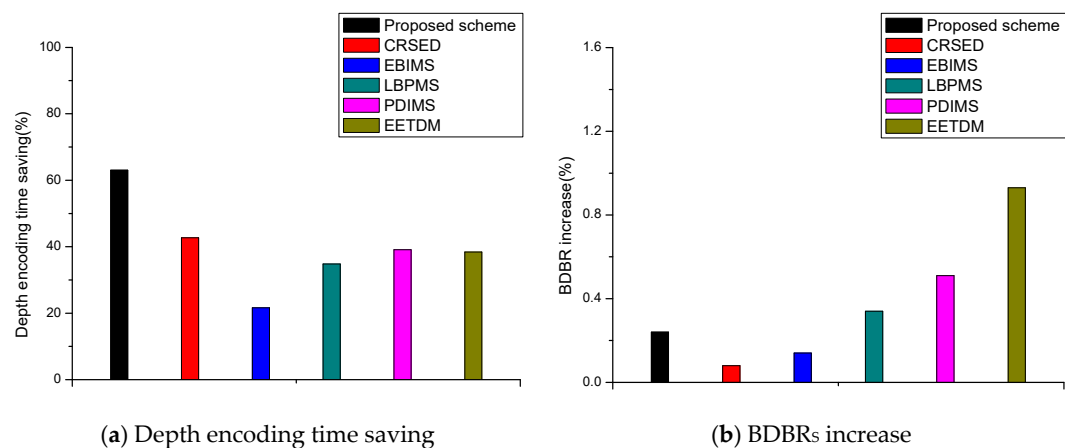


Figure 7. Coding results of the proposed scheme and five algorithms under the “All-Intra” configuration.

5. Conclusions

In this paper, we presented a fast scheme to reduce the complexity of depth coding, which comprises two strategies: early SKIP mode decision and adaptive intra prediction selection. The basic idea of this scheme is to use the spatial–temporal and texture video correlations to analyze current treeblock characteristics and skip some unnecessary modes. The experiment results demonstrate that the proposed scheme can achieve the computational saving load by 63.0% compared with the HTM 16.1, with only a 0.24% BDBR increase for the rendered video. Furthermore, the proposed scheme outperforms well-known fast methods with the better complexity reduction.

Author Contributions: Conceptualization, Q.Z. and Y.W.; Formal Analysis, Q.Z. and B.J.; Funding acquisition, Q.Z. and Y.G.; Methodology, Q.Z. and B.J.; Resources, Y.W. and T.W.; Software, Y.W. and B.J.; Supervision, Y.G.; Validation, Q.Z. and B.J.; Writing—Original Draft, Q.Z.; Writing—Review & Editing, Q.Z. and Y.W.

Funding: This work was supported in part by the National Natural Science Foundation of China (Grant No. 61771432, 61302118, and 61702464), Scientific Project (Grant No. 182102210156, and 182102210610), Innovation Talents (Grant No.17HASTIT022), and the Education Department Project (Grant No.18B510019, and 17B510011).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tanimoto, M.; Tehrani, M.P.; Fujii, T.; Yendo, T. Free-Viewpoint TV. *IEEE Signal Process. Mag.* **2011**, *28*, 67–76. [[CrossRef](#)]
2. Müller, K.; Merkle, P.; Wiegand, T. 3-D video representation using depth maps. *Proc. IEEE.* **2011**, *99*, 643–656. [[CrossRef](#)]
3. Chen, Y.; Vetro, A. Next-Generation 3D Formats with Depth Map Support. *IEEE MultiMed.* **2014**, *21*, 90–94.
4. Tech, G.; Chen, Y.; Müller, K.; Ohm, J.; Vetro, A. Overview of the multiview and 3D extensions of High Efficiency Video Coding. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *26*, 35–49. [[CrossRef](#)]
5. Fehn, C. Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. In Proceedings of the SPIE Stereoscopic Displays and Virtual Reality Systems XI, San Jose, CA, USA, 19–21 January 2004; pp. 93–104.
6. Sullivan, G.J.; Boyce, J.M.; Ying, C.; Ohm, J.-R.; Segall, C.A.; Vetro, A. Standardized Extensions of High Efficiency Video Coding (HEVC). *IEEE J. Sel. Top. Sign. Process.* **2013**, *7*, 1001–1016. [[CrossRef](#)]
7. Sullivan, G.J.; Ohm, J.-R.; Han, W.-J.; Wiegand, T. Overview of the high efficiency video coding (HEVC) standard. *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *22*, 1649–1668. [[CrossRef](#)]
8. Müller, K.; Schwarz, H.; Marpe, D.; Bartnik, C.; Bosse, S.; Brust, H.; Hinz, T.; Lakshman, H.; Merkle, P.; Rhee, H.; et al. 3D high efficiency video coding for multi-view video and depth data. *IEEE Trans. Circuits Syst. Video Technol.* **2013**, *22*, 3366–3378. [[CrossRef](#)]
9. Chen, Y.; Tech, G.; Wegner, K.; Yea, S. Test Model 11 of 3D-HEVC and MV-HEVC. In Proceedings of the MPEG 111, Geneva, Switzerland, 16–20 February 2015. Document JCT3V-J1003.
10. Yoon, D.H.; Ho, Y.S. Fast mode decision algorithm for depth coding in 3-D video systems using H.264/AVC. In Proceedings of the 5th Pacific-Rim Symposium on Image and Video Technology, Gwangju, South Korea, 20–23 November 2011; pp. 25–35.
11. Lee, J.Y.; Wey, H.-C.; Park, D.-S. A fast and efficient multi-view depth image coding method based on temporal and inter-view correlations of texture images. *IEEE Trans. Circuits Syst. Video Technol.* **2011**, *2*, 1859–1868. [[CrossRef](#)]
12. Shen, L.; Zhang, Z.; Liu, Z. Inter mode selection for depth map coding in 3D video. *IEEE Trans. Consum. Electron.* **2012**, *58*, 926–931.
13. Yoon, D.; Ho, Y. Fast depth video coding method using adaptive edge classification. *Circuits Syst. Signal Process.* **2013**, *32*, 803–813. [[CrossRef](#)]
14. Cernigliaro, G.; Jaureguizar, F.; Cabrera, J.; García, N. Low complexity mode decision and motion estimation for H.264/AVC based depth maps encoding in free viewpoint video. *IEEE Trans. Circuits Syst. Video Technol.* **2013**, *23*, 769–783. [[CrossRef](#)]
15. Pan, Z.; Zhang, Y.; Kwong, S. Fast mode decision based on texture–depth correlation and motion prediction for multiview depth video coding. *J. Real Time Image Process.* **2016**, *11*, 27–36. [[CrossRef](#)]
16. Shen, L.; Zhang, Z. Efficient depth coding in 3D video to minimize coding bitrate and complexity. *Multimed. Tools Appl.* **2014**, *72*, 1639–1652. [[CrossRef](#)]
17. Zhang, Q.; An, P.; Zhang, Y.; Shen, L.; Zhang, Z. Low complexity multiview video plus depth coding. *IEEE Trans. Consum. Electron.* **2011**, *57*, 1857–1865. [[CrossRef](#)]
18. Tohidypour, H.R.; Pourazad, M.T.; Nasiopoulos, P.; Leung, V. A content adaptive complexity reduction scheme for HEVC-based 3D video coding. In Proceedings of the 18th International Conference on Digital Signal Processing (DSP), Santorini, Greece, 1–3 July 2013; pp. 1–5.
19. Shen, L.; An, P.; Zhang, Z.; Hu, Q.; Chen, Z. A 3D-HEVC fast mode decision algorithm for real-time applications. *ACM Trans. Multimed. Comput. Commun. Appl.* **2015**, *11*, 1–23. [[CrossRef](#)]

20. Zhang, Q.; Chen, M.; Huang, X.; Li, N.; Gan, Y. Low-complexity depth map compression in HEVC-based 3D video coding. *EURASIP J. Image Video Process.* **2015**, *2015*, 2. [[CrossRef](#)]
21. Conceição, R.; Avila, G.; Corrêa, G.; Porto, M.; Zatt, B.; Agostini, L. Complexity reduction for 3D-HEVC depth map coding based on early Skip and early DIS scheme. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 1116–1120.
22. Chen, H.; Fu, C.; Zhang, Y.; Chan, Y.; Siu, W. Early merge mode decision for depth maps in 3D-HEVC. In Proceedings of the 22nd International Conference on Digital Signal Processing (DSP), London, UK, 23–25 August 2017; pp. 1–5.
23. Lei, J.; Duan, J.; Wu, F.; Ling, N.; Hou, C. Fast mode decision based on grayscale similarity and inter-view correlation for depth map coding in 3D-HEVC. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *28*, 706–718. [[CrossRef](#)]
24. Chung, K.; Huang, Y.; Lin, C.; Fang, J. Novel bitrate saving and fast coding for depth videos in 3D-HEVC. *IEEE Trans. Circuits Syst. Video Technol.* **2016**, *26*, 1859–1869. [[CrossRef](#)]
25. Zhang, Q.; Zhang, N.; Wei, T.; Huang, K.; Qian, X.; Gan, Y. Fast depth map mode decision based on depth-texture correlation and edge classification for 3D-HEVC. *J. Visual Commun. Image Represent.* **2017**, *45*, 170–180. [[CrossRef](#)]
26. Zhang, M.; Zhao, C.; Xu, J.; Bai, H. A fast depth-map wedgelet partitioning scheme for intra prediction in 3D video coding. In Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS), Beijing, China, 19–23 May 2013; pp. 2852–2855.
27. Sanchez, G.; Saldanha, M.; Balota, G.; Zatt, B.; Porto, M.; Agostini, L. Complexity reduction for 3D-HEVC depth maps intra-frame prediction using simplified edge detector algorithm. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 3209–3213.
28. Park, C. Efficient intra-mode decision algorithm skipping unnecessary depth-modelling modes in 3D-HEVC. *Electron. Lett.* **2014**, *51*, 756–758. [[CrossRef](#)]
29. Park, C. Edge-Based Intramode Selection for Depth-Map Coding in 3D-HEVC. *IEEE Trans. Image Process.* **2015**, *24*, 155–162. [[CrossRef](#)]
30. Gu, Z.; Zheng, J.; Ling, N.; Zhang, P. Low complexity Bi-Partition mode selection for 3D video depth intra coding. *Displays* **2015**, *40*, 2–8. [[CrossRef](#)]
31. He, G.; Hu, J.; Li, Y.; Yu, W.; Yang, Z.; Liu, P.; Guo, R. Fast mode decision and PU size decision algorithm for intra depth coding in 3D-HEVC. *J. Visual Commun. Image Represent.* **2017**, *49*, 303–314. [[CrossRef](#)]
32. Zhang, H.; Fu, C.; Chan, Y.; Tsang, S.; Siu, W. Probability based depth intra mode skipping strategy and novel VSO metric for DMM decision in 3D-HEVC. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *28*, 513–527. [[CrossRef](#)]
33. Zhang, Q.; Li, N.; Huang, L.; Gan, Y. Effective early termination algorithm for depth map intra coding in 3D-HEVC. *Electron. Lett.* **2014**, *50*, 994–996. [[CrossRef](#)]
34. Zhang, Q.; Yang, Y.; Chang, H.; Zhang, W.; Gan, Y. Fast intra mode decision for depth coding in 3D-HEVC. *Multidimens. Syst. Signal Process.* **2017**, *28*, 1203–1226. [[CrossRef](#)]
35. 3D-HEVC software HTM-16.1. Available online: https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-16.1/ (accessed on 9 April 2019).
36. Mueller, K.; Vetro, A. Common test conditions of 3DV core experiments. In Proceedings of the MPEG 107, San Jose, CA, USA, 13–17 January 2014. Document JCT3V-G1100.
37. Shen, L.; Zhang, Z.; Liu, Z. Effective CU size decision for HEVC intracoding. *IEEE Trans. Image Process.* **2014**, *23*, 4232–4241. [[CrossRef](#)]
38. Tanimoto, M.; Fujii, T.; Suzuki, K. View Synthesis algorithm in view synthesis reference software 2.0 (VSRS 2.0). In Proceedings of the ISO/IEC JTC 1/SC 29/WG 11 Meetings, Lausanne, Switzerland, 2–6 February 2009. Document M16090.
39. Bjontegaard, G. Calculation of Average PSNR differences between RD curves. In Proceedings of the VCEG Meeting, VCEG-M33, Austin, TX, USA, 2–4 April 2001. Document ITU-T SG16 Q.6.

