Journal of Information Hiding and Multimedia Signal Processing ©2018 ISSN 2073-4212
Ubiquitous International Volume **9**, Number **5**, September **2018**

Fast Mode Decision Algorithm for Coding Depth Maps in 3D High-Efficiency Video Coding

HaiWei Lei*, PinLe Qin and YanBo Wang

School of Computer Science and Technology
North University of China
No.3, University Road, Taiyuan, 030051, Shanxi, China

AnHong Wang

Institute of Digital Multimedia and Communication
Taiyuan University of Science and Technology
No. 66, Waliu Road, Taiyuan, 030024, China
*Corresponding Author:lhw0312@163.com

Received February, 2018; revised May, 2018

Abstract. *The mode decision is a computationally expensive process in the three-dimensional extension of high-efficiency video coding (3D-HEVC), which is particularly prominent when coding depth maps. A fast algorithm is proposed that includes three improvements to reduce its computational complexity. In the mode decision process, a coding unit (CU) can be encoded directly or split into four smaller CUs. If the rate distortion (RD) cost of this CU is less than the total RD cost of its four sub-CUs, this CU is directly encoded. Otherwise, it will be split into four smaller sub-CUs. We propose to terminate the CU partition as soon as the aggregated RD cost of the first k ($k < 4$) sub-CUs exceeds that of their parent CU instead of completing the calculation of all sub-CUs. In addition, based on the fact that the newly-added depth intra skip (DIS) mode does not encode residual data, we propose two other improvements, namely the early determination of the optimal prediction unit (PU) mode by evaluating the RD cost of the DIS mode and the termination of the CU partition if the best PU mode is DIS mode and its RD cost equals 0. The experimental results showed that the proposed algorithm achieved average encoding time savings of 5.92% and 31.12% for the common test conditions and the all-intra configuration, respectively, with negligible loss of coding performance.*
Keywords: Mode decision, Depth map coding, CU partition, Early termination algorithm

---

1. **Introduction.** The three-dimensional extension of high-efficiency video coding (3D-HEVC) is the state-of-the-art video coding standard for efficient compression of 3D video with the multi-view video plus depth (MVD) format [1]. In the MVD format, a small number of videos captured from different viewpoints and the associated depth maps are encoded at the encoder side. At the decoder side, after the video and depth data have been decoded, additional intermediate views can be synthesized using depth-image-based rendering (DIBR) [2] technique. As one of the applications, these views are suitable for displaying on an autostereoscopic display, and a pair of them can be seen by the audience according to different viewing positions and distances. In this way, people can enjoy 3D videos without the need for special glasses or other head gear.

The HEVC encoder has high computational complexity, with one of the primary factors being the mode decision. In contrast to the macroblock of its predecessors standards, the analogous structure in HEVC is the coding tree unit (CTU), which also is referred to as the largest coding unit. When coding is initiated, each image frame is split into many equal-sized CTUs. Then, these CTUs go through a mode decision process. A CTU is split recursively into smaller coding units (CUs). For each CU, the mode decision algorithm will test all intra and inter prediction unit (PU, the basic unit for making predictions.) modes to determine an optimal PU mode. In the test process, an appropriate intra-prediction mode for each intra PU must be selected, or a motion vector (MV) must be derived for each inter PU. In addition, the transform unit (TU), which is the basic unit for transforming and quantizing, also must be determined in the mode decision process. This complex process is performed by means of a special data structure that is referred to as quadtree.

The problem of high computational complexity is more prominent in the 3D-HEVC encoder. Because the encoder must simultaneously encode at least two texture videos and the associated depth maps, the processing required is at least four times greater than that required of the HEVC encoder. Furthermore, since depth maps have features that are different from those of the texture video, i.e., features that are characterized by slowly varying (or flat) areas with sharp edges, the use of the original coding tools in HEVC usually produces artifacts in the edge regions. Therefore, a modified version of HEVC that incorporates new coding tools, such as the depth-modeling mode (DMM) [3], depth intra skip (DIS) [4], and simplified depth coding (SDC) [5], is used to better encode the depth maps. These coding tools improve the coding performance, but they also take an unacceptable amount of time. As a result, the computational complexity of the coding of depth maps is much greater than that of coding texture videos.

Also, the complexity for encoding depth maps is concentrated mainly in the mode decision. The depth map mode decision still is based on the quadtree structure, which has the inherent characteristic of high computational complexity. At the same time, the use of new coding tools increases the number of candidate prediction modes, which results in a large number of additional computations. When SDC is used as an alternative method for residual coding of the depth map, the problem is exacerbated, and the decision concerning the depth map mode becomes surprisingly complicated. Therefore, it is desirable to design an algorithm that can provide a fast depth map mode decision to reduce the computational complexity.

In this paper, we analyzed the depth map mode decision in 3D-HEVC and developed a fast mode decision algorithm for coding depth maps. By comparing the aggregated RD cost of the first k ($k < 4$) sub-CUs with that of their parent CU, the algorithm determines whether the calculation of the RD cost of the tail sub-CUs should be skipped. Moreover, based on the characteristic that the DIS mode does not encode the residual data after the prediction, our algorithm early decides whether a CU should be further split into smaller CUs by evaluating the RD cost of the DIS mode and simplifies the mode selection process by skipping the testing of some PU modes. With these measures, the proposed algorithm skips a part of the complex calculations, thereby speeding up the process of the depth map mode decision.

The remainder of this paper is organized as follows. In Section 2, we provide an overview of related work on fast mode decision algorithms for the 3D-HEVC encoder. Section 3 details the computational complexity associated with the mode decision. In Section 4, we propose a fast depth map mode decision algorithm. Our experimental results are presented in Section 5, and our conclusions are given in Section 6.

2. **Related work.** Several fast mode decision algorithms have been proposed to reduce the computational complexity of the HEVC mode decision. These algorithms can be

classified into three categories, i.e., straightforward methods, spatiotemporal correlation-based methods, and methods based on advanced algorithms.

The straightforward methods tend to omit unnecessary calculations according to the results derived from previous steps. The method in [6] reduced the complexity of the mode decision by checking the coded block flag (CBF); if the CBF of an inter PU mode for the luma and two chromas are both zero, the process of testing the remaining PU modes can be skipped. In [7], an enhanced SKIP condition, the differential motion vector (DMV) and the CBF of the inter 2N×2N PU modes, respectively, were equal to (0, 0) and zero was used to skip the remaining PU modes. Different from [6], this condition is checked only once at the beginning of the mode decision. A tree-pruning algorithm was proposed [8] to skip the sub-tree computations when the best prediction mode of the current CU node is the SKIP mode. Due to their simple and ease of implementation, these three methods were used in the HEVC encoder. In addition, various user-defined conditions have been proposed to determine whether to early terminate CU partition. In [9], the mean square error (MSE) of the prediction residual block was used as a basis for terminating CU partition. The authors of [10] proposed to terminate the current CU partition when its RD cost was below a certain threshold, whereas [11] determined whether to terminate CU partition by evaluating the Sobel edge density of the coding tree block. In [12], a fast CU depth decision algorithm was proposed for the inter prediction of HEVC that terminated the CU partition early according to the motion characteristics (the cost of motion vector) of the CUs.

Methods in the second category sped up the mode decision by reducing the depth search range of CUs; it mainly utilized the depth information of neighboring CUs or the texture characteristics of the CU itself. In [13, 14], fast algorithms for making CU depth decisions based on spatiotemporal correlations were proposed by considering the fact that the depth of the current CTU has a significant correlation with that of its neighboring CTUs, where the depth of the current CTU can be predicted from both its spatial and temporal neighboring CTUs. The method in [15] reduced the range of the depth of the CUs by skipping some specific depth levels that rarely are used in the previous frame and neighboring CUs. A similar method was proposed in [16], and it jointly uses the inter-level correlation of the quadtree structure and the spatiotemporal correlation to adaptively choose the inter mode for HEVC. An early MERGE mode decision algorithm was proposed [17] based on the statistical results, i.e., if the root CU is encoded with the MERGE mode, its children CUs also have a large probability to be encoded with the MERGE mode. In [18], the CU depth pruning strategies were determined adaptively according to the standard deviation of the statistical spatiotemporal depth information. Methods for controlling the complexity of HEVC were proposed based on a decision algorithm that dynamically adjusts the depth range of the CUs [19].

Methods in the third category improved the mode decision by using advanced algorithms. Bayesian decision rules were used in [20-22], and the support vector machine (SVM) was used in [23-25] to help to determine the size of the CU. Also, [26] used pyramid motion divergence (PMD) to help select the size of CU. In addition, a scheme for allocating the hierarchical complexity [27] was developed that used linear programming to allocate the computational complexities among the frames.

All of the algorithms mentioned above are well developed for the HEVC mode decision, and they save significant encoding time and provide acceptable coding performance. However, these fast algorithms are ineffective or not directly applicable to the new 3D-HEVC standard. Recently, studies also have been reported on the reduction of the computational complexity of the 3D-HEVC mode decision with the aim of speeding up the coding of the depth map. Some of them borrowed from the fast mode decision methods used in HEVC.

The others mainly cover two categories, i.e., 1) methods based on co-located texture video information and 2) methods based on DMM.

In 3D-HEVC, the texture video is coded prior to its associated depth map, and the texture video and the depth map represent the same scene at the same moment. Therefore, the coding information of texture videos can provide a reference in coding the depth map. A quadtree limitation algorithm with a predictive coding of the quadtree was proposed for coding depth maps in [28], where a given depth CU cannot be split more than its co-located CU in the texture video. In [29], the motion vectors and prediction mode drawn from the texture video were used to accelerate the procedure for making the mode decision of depth maps. Another fast algorithm reduced the computational complexity of the mode decision by skipping some specific depth intra prediction modes that rarely are used in co-located texture CUs [30]. An inter-component (texture-depth) predictor based on the similarities between texture videos and depth maps was designed in [31] to simplify the mode decision of the depth map.

DMM was developed to provide a better representation of the edges of depth maps, and it is not suitable for coding blocks in flat (or gradually changing) regions. For flat coding blocks, DMM must not be added to the full-RD search list for the full-RD cost calculation. With this consideration, many fast algorithms have been proposed. A fast DMM selection algorithm was proposed in [32], based on Most Probable Mode (MPM). In [33], the authors proposed to skip DMM when the variance of the current CU is lower than a given threshold. In [34], a scheme with a simplified edge detector (SED) and a gradient-based mode one filter (GMOF) was used to determine whether DMM should be skipped. In [35], a fast algorithm based on the edge classification in the Hadamard transform domain was developed to make the mode decision faster. The method in [36] used the calculated RD costs of the conventional intra-modes to skip the unnecessary depth-modeling modes, thereby reducing the computational complexity of the mode decision of the depth map.

With the introduction of the latest prediction mode, new methods based on mode own characteristics have been developed. In [37], a tree pruning strategy based on the single depth intra (SDI) mode was used to reduce the computational complexity of the depth map intra-mode decision. The algorithm in this paper is this kind of method, and it is based on the newly-added DIS mode. It is a novel method to reduce the computational complexity associated with making the mode decision.

3. **Computational complexities in the mode decision of the depth map.** The overall mode decision algorithm consists of an optimal CU partition decision and an optimal PU mode decision (or PU mode selection). That is, the CU partition decision must derive the optimal CU partition of a CTU using a quadtree, where the sizes of the CUs in different positions are determined uniquely. At the CU level, the PU mode decision must select an optimal PU mode for each CU by testing all inter and intra PU modes. In addition, new coding tools for depth maps are involved in the mode decision process. Therefore, this section will describe the three aspects of the computational complexity of the mode decision.

3.1. **Recursive CU Partitioning.** CU partitioning is performed by using a hierarchical quadtree. Fig. 1 shows that a complete quadtree consists of 85 CU nodes when its depth level is four. The encoder traverses all of these nodes in depth-first order, and the numbers in Fig. 1 indicate this access order.

When visiting a certain CU node, the encoder typically investigates all inter and intra PU modes by computing their RD cost, and then it selects the optimal PU mode (i.e., the mode that minimizes the RD cost) for this CU node. A CU can be encoded directly or

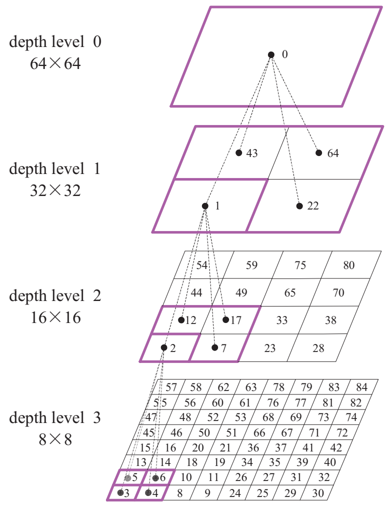

FIGURE 1. Process of CU quadtree partitioning with a depth level of four

split into four smaller CUs with each of them encoded separately. This decision depends on the result of the comparison between the RD cost of this CU and the total RD cost of its four sub-CUs. That is, if the following inequality holds, the current CU is directly encoded. Otherwise, it will be split into four smaller CUs.

$$J_{CU_i^d}(m_i) < \sum_{j=0}^{3} J_{CU_{i+\alpha\cdot j+1}^{d+1}}(m_{i+\alpha\cdot j+1}) \tag{1}$$

where d is the depth level, m is the optimal PU mode for a CU, α is 21, 5 and 1 respectively, when d equals 0, 1 and 2. Note that the comparison is performed layer by layer and from bottom to top. The partition instance with the smaller RD cost is selected and regarded as a substitute for the current CU node; then, it is compared with its parent CU node together with the other three sibling CU nodes, which are assumed to have completed the comparison with their sub-CU nodes. Finally, the algorithm provides the optimal CU partition. Fig. 2 illustrates a result of CU partition, including the final situation of CU partition and the corresponding quadtree. The above procedure shows that the complexity of CU partition is quite high because the algorithm must traverse all of the CU nodes.

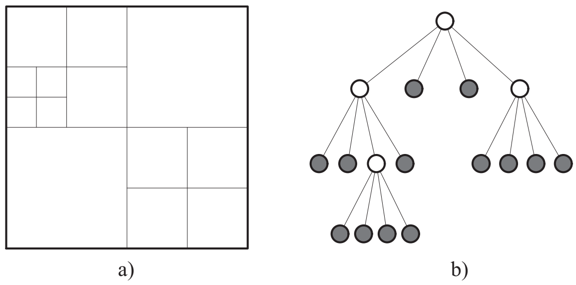

FIGURE 2. Example of CU partitioning: a) the resulting CU partition; b) the corresponding quadtree

3.2. **Exhaustive Method for Selecting the PU Mode.** The selection of the PU mode implicitly is performed in the process of CU partitioning. It is responsible for identifying the optimal PU mode for a given CU node in the quadtree by testing all inter and intra

PU modes, as shown in Fig. 3. When testing a certain PU mode, the encoder must find the best intra prediction mode for an intra PU or derive a motion vector for an inter PU. This means that a large number of calculations must be made in each PU mode. Taking a 2N×2N intra mode as an example, the DC, Planar, and 33 other directional modes will be tested exhaustively with the aim of determining the best intra prediction mode for this 2N×2N intra PU. Furthermore, partitioning of the transform unit (TU) is performed simultaneously in the process of selecting the PU mode; it is responsible for seeking an appropriate transform block size to obtain the best transform and quantization results. It can be seen that the selection of the PU mode for one CU node is computationally expensive, and this expensive process must be executed repeatedly for each CU node in the quadtree.

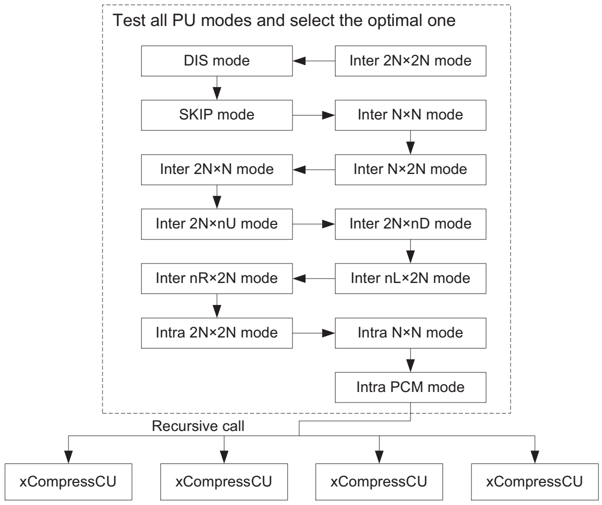

Figure 3. Process of selecting PU modes

3.3. **Introduction of New Coding Tools.** New coding tools, including DIS, DMM, and SDC, are involved in the mode decision of the depth map, and this increases the computational complexity of the mode decision.

SDI is a new, intra coding mode just for coding depth maps. When a CU is coded as the SDI mode, the prediction is performed by filling the CU with a single depth value, and the residual is not encoded. That is, the SDI mode fills a CU with one sample from the spatial neighboring pixels for the reconstruction. Unlike SDI, the DIS mode uses the vertical and horizontal prediction modes for the reconstruction instead of using a spatial neighboring pixel. In the current 3D-HEVC reference software, SDI and DIS are harmonized and collectively referred to as the DIS mode. So, the DIS mode consists of four types of coding modes. In the process of selecting the PU mode, the encoder will test all four of these coding modes, determine the best one, and compare it with the other PU modes.

DMM is a model-based intra coding mode. In DMM, a depth block is approximated by a model that partitions the area of the block into two non-rectangular regions, with each region represented by a constant value. DMM consists of two steps, i.e., 1) region partitioning and 2) the search of the optimal constant partition values (CPVs) for the resulting regions. Unfortunately, both of these steps are expensive computationally. For example, in DMM1 (Explicit Wedgelet signaling), the encoder must test a large number of partition patterns (even using the fast search mode) [38] and determine the best one through the RD optimization. The numbers of partition patterns for different sizes of blocks are listed in Table 1. After determining the partition pattern, the encoder still must use a search algorithm to search all possible combinations of the two CPVs to determine the optimal CPV pair for these two regions.

TABLE 1. Number of Wedgelet patterns tested for the full and fast search

block size	full search	fast search
4×4	86	$\leq 58 + 8$
8×8	782	$\leq 310 + 8$
16×16	1394	$\leq 338 + 8$
32×32	1503	$\leq 368 + 8$

SDC is an alternative method for residual coding of depth maps, and it is different from the conventional transform-based residual coding approach. In SDC, only a DC residual is coded for each partition within a PU, and transform and quantization are skipped. All intra prediction modes (excluding the planar mode) and DMM can work with the traditional HEVC and SDC residual coding methods. That is, the encoder tests both the traditional HEVC and SDC residual coding methods for a 2N×2N intra PU, and then it selects the better one. This doubles the complexity of the coding.

4. **Proposed Fast Mode Decision Algorithm.** Although the 3D-HEVC mode decision algorithm is well designed, and many other methods have been presented, there is still some room for improvement in the decision process associated with the 3D-HEVC depth map coding. By comparing Fig. 1 and Fig. 2, it can be found that a number of CU nodes need not to be split further. The partition can be avoided if we know these CU nodes in advance. In addition, the optimal PU mode may be determined earlier by testing a part of the PU modes rather than testing all of them in the process of selecting the PU mode. Based on the above considerations, we propose three key points to improve the mode decision of the depth map.

4.1. **Skipping tail sub-CUs using the aggregated RD cost of partial sub-CUs.** As described in the previous section, whether a CU is directly encoded or split into four smaller CUs depends on the result of the comparison between the RD cost of this CU and the total RD cost of its four sub-CUs. To make the comparison, the RD cost of the current CU and its four sub-CUs must be calculated in advance. In the existing algorithm, the RD costs of all four sub-CUs and their sum are calculated. It is apparent that the calculation of RD cost for the remaining sub-CUs is unnecessary if the sum of RD cost of the first k (k < 4) sub-CUs has exceeded the RD cost of the current CU. Therefore, we proposed to skip the calculation of RD cost of the remaining sub-CU nodes when inequality (2) holds.

$$J_{CU_i^d} < \sum_{j=0}^{k} J_{CU_{i+\alpha\cdot j+1}^{d+1}} \tag{2}$$

Note that the skipped node may be a leaf node or a subtree, depending on the depth level at which the skip operation takes place. Fig. 4 shows the case in which two leaf nodes and a subtree are pruned, meaning that the RD cost of these nodes will no longer be calculated.

A similar approach was proposed in [39] for improving the HEVC mode decision, and this approach used the predicted RD cost as an approximation of the actual RD cost of

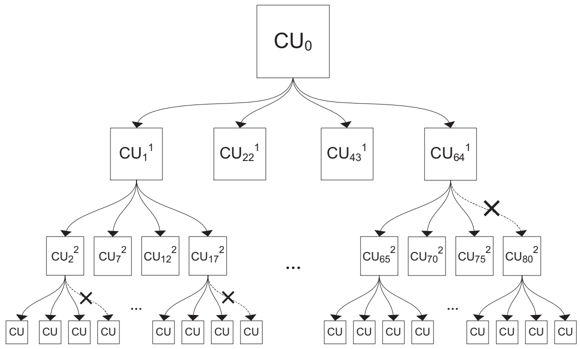

FIGURE 4. Skipping tail sub-CUs using aggregated RD cost of partial sub-CUs

four sub-CUs, and Eq. (3) is used to calculate the predicted RD cost.

$$\hat{J}^{RD}_{CU^{d+1}} = \left(min \left\{ \frac{4}{k}, \frac{J^{H}_{CU^{d+1}}}{J^{H}_{CU^{d+1},k}} \right\} \right) \cdot J^{RD}_{CU^{d+1},k} \tag{3}$$

where $J^{H}_{CU^{d+1}}$ is the total Hadamard cost for all four sub-CUs and $J^{H}_{CU^{d+1},k}$ and $J^{RD}_{CU^{d+1},k}$ are the aggregated Hadamard RD cost and aggregated actual RD cost of the first k sub-CUs, respectively. Then, after adding a scaling factor, β, the inequality (4) is regarded as the criterion for determining whether or not to skip the remaining sub-CU nodes.

$$\hat{J}^{RD}_{CU^{d+1}} > \beta_k \cdot J^{RD}_{CU^{d}} \tag{4}$$

We call inequality (4) the variant of inequality (2). To determine the truth, the case in which inequality (2) is applied directly to the HEVC intra mode decision was tested, and the results are provided in Table 2. It can be see that the average encoding time saving (ΔT) was only 2.64%. This is probably why [39] uses inequality (4) instead of directly using inequality (2). However, the effect of the direct application of inequality (2) to the 3D-HEVC intra mode decision is obvious. Table 3 shows the results, and they show a significant savings of encoding time, with an average ΔT of about 25%. This implies that the complexity of the 3D-HEVC intra mode decision is reduced by using inequality (2) to skip the tail sub-CU nodes.

TABLE 2. Total savings of encoding time when inequality (2) is applied to HEVC

Class	Sequence	Resolution	$\Delta T(\%)$
A	Traffic	2560×1600	1.93
B	BasketballDrive	1920×1080	1.66
C	BasketballDrill	832×480	0.25
D	BasketballPass	416×240	1.80
E	FourPeople	1280×720	4.73
F	ChinaSpeed	1024×768	5.44
Average			2.64

4.2. **Early determination of the optimal PU mode by evaluating the distortion of the DIS mode.** It was observed that, with the introduction of the DIS mode, the optimal PU mode can be determined by testing some of the PU modes rather than testing all of them in the PU mode selection process.

TABLE 3. Total savings of encoding time when inequality (2) is applied to 3D-HEVC

Sequence	Resolution	$\Delta T(\%)$
Balloons	1024×768	20.60
Kendo	1024×768	22.91
Newspaper_CC	1024×768	15.95
GT_Fly	1920×1088	32.42
Poznan_Hall2	1920×1088	29.55
Poznan_Street	1920×1088	29.11
Undo_Dancer	1920×1088	26.12
Shark	1920×1088	25.46
Average		25.27

The selection of the optimal PU mode is performed by means of RD optimization. The RD cost of every available PU mode can be calculated by Eq. (5), and the PU mode that minimizes J is selected.

$$J = D(m) + \lambda \cdot R(m) \tag{5}$$

where λ is the Lagrange multiplier, D(m) is the distortion of the current CU and is computed as a weighted average of the synthesized view distortion and the depth map distortion; and R(m) is the number of bits required for encoding this CU, which consists of two parts as follows:

$$R(m) = R_P(m) + R_R(m) \tag{6}$$

where $R_P(m)$ and $R_R(m)$ are the numbers of bits required for encoding the prediction parameters (PPs), such as the inter/intra PU partition, and the residual data of the current CU.

In the DIS mode, only two PPs, i.e., DISFlag and DISType, are required to signal whether the DIS mode is used and the sub-type that is used. The residual is not encoded after the prediction, which means $R_R(m) = 0$. In contrast, other PU modes must signal at least two parameters and the potential residual data. Therefore, the DIS mode has advantages compared to other PU modes in terms of R(m). RD cost will be very small when the DIS mode is used and the resulting distortion is sufficiently small (such as zero). In this case, there is a high probability that the DIS mode will be selected as the optimal PU mode.

Note that the number of bits required for coding PPs is related to the number of PPs and to the coding technique that is used. In 3D-HEVC, the encoder uses context-based adaptive binary arithmetic coding (CABAC) as the only entropy coding tool. The "context" is a probability model for coding a symbol, and it can be chosen from a selection of available models. The model that is chosen depends on the statistics of the recently-coded data symbols. That is, the number of bits encoded by the CABAC for the same symbol can be more or less, depending on its context. Therefore, further investigation is required to support the above, less-stringent ideas. The first 10 frames of test sequences are encoded using different quantization parameters (QPs) at 25, 30, 35, 40, and the percentages of the cases in which the DIS mode is selected as the best PU mode when the resulting distortion equals zero are counted. Table 4 shows the results, it is apparent that the DIS mode wins with a high percentage.

Based on the analysis above, an approach was achieved for the early determination of the optimal PU mode. Specifically, first, a parameter called "distItem" is added to the

TABLE 4. Percentage of cases in which the DIS mode was selected as the best PU mode when the distortion was zero

Sequence	QP25(%)	QP30(%)	QP35(%)	QP40(%)
Balloons	99.38	99.61	99.79	99.83
Kendo	99.63	99.72	99.86	99.91
Newspaper_CC	98.63	99.31	99.56	99.68
GT_Fly	99.62	99.97	99.96	100
Poznan_Hall2	99.84	99.94	99.96	99.98
Poznan_Street	98.06	99.85	99.97	100
Undo_Dancer	99.43	99.86	99.93	99.96
Shark	99.4	99.29	99.64	99.95
Average	99.25	99.69	99.83	99.91

parameter list of DIS function to record the distortion. In addition, a distItem-based decision criterion is added after the DIS mode has been tested. If distItem = 0, the DIS mode is regarded as the optimal PU mode, and the PU mode selection process is terminated early. Fig. 6 shows this process, with the flow path of the mode decision proceeding along the left branch at the first green decision symbol, thereby avoiding the need to investigate the remaining PU modes.

4.3. **Skipping all sub-CUs using the DIS mode.** In Section 4.1, inequality (2) is used to skip the tail sub-CU nodes. However, more sub-CU nodes could be skipped after the DIS mode is added as one of candidate PU modes.

According to Eq. 5, $R_R(m)$ will be 0 and the RD cost will depend directly on $R_P(m)$ when D(m) = 0. At this point, inequality (1) is simplified as:

$$R_{P_{CU_i^d}}(m_i) < \sum_{j=0}^{3} R_{P_{CU_{i+\alpha\cdot j+1}^{d+1}}}(m_{i+\alpha\cdot j+1}) \tag{7}$$

In all probability, the number of bits required for encoding the PPs of one CU is less than that required for four CUs. This possibility increases considerably when the current CU is predicted by the DIS mode, because only two PPs are required for signaling. Thus, another criterion was added to our algorithm before CU partitioning. That is, if the DIS mode is selected as the optimal PU mode of the current CU node and distItem = 0, this CU is no longer split into four smaller sub-CUs, which means that the local optimal CU partition is determined directly without computing the RD cost of sub-CU nodes. It actually uses the idea of the branch and bound algorithm. Fig. 5 shows an example with three places skipped, indicating that these nodes no longer go through the PU mode selection process.

Note that the method described here is different from that in Section 4.1. In this method, all four sub-CU nodes are pruned together without calculating the RD cost of any one of them. In Section 4.1, only the tail sub-CU nodes are skipped, and the RD cost of the first k (k < 4) sub-CUs still must be calculated. In spite of this, these two methods are complementary.

4.4. **Overall Algorithm.** The proposed overall algorithm incorporates the above three key points to reduce the computational complexity of the mode decision. A flowchart of the improved mode decision process is shown in Fig. 6, where the decisions in yellow were added in [7, 8], and those in green were added by our algorithm. Note that the data required in the decision symbol already have been calculated in the original 3D-HEVC

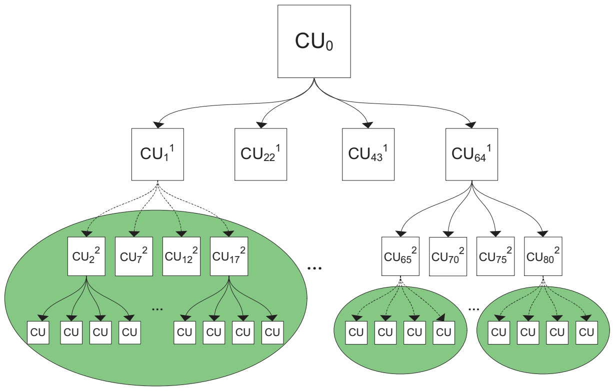

Figure 5. Skipping all sub-CUs using DIS mode

encoder. Therefore, our algorithm does not cause additional computational overhead, and it is easy to implement.

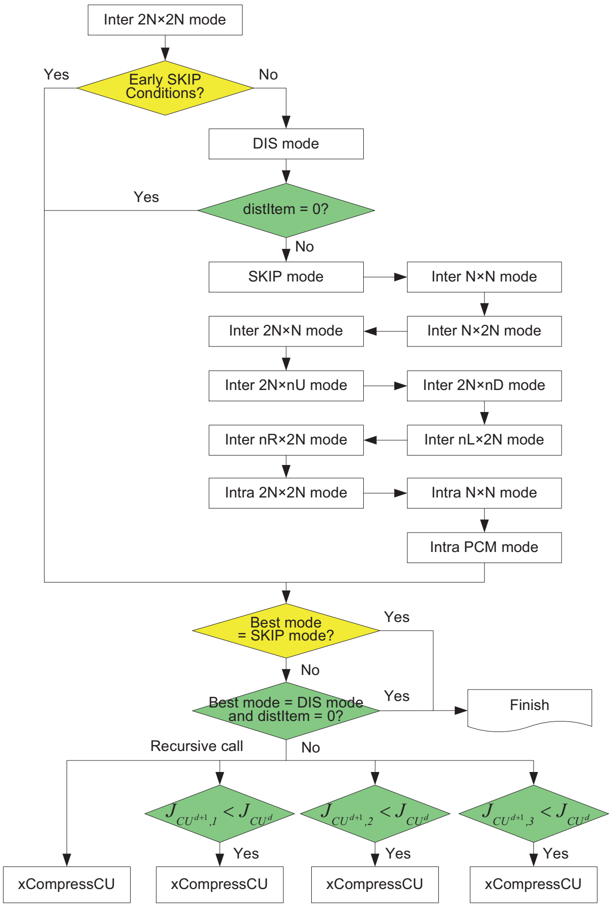

Figure 6. Flowchart of the improved depth map mode decision

5. **Experimental results and analysis.** To demonstrate the efficiency of the proposed fast mode decision algorithm, we performed experiments on HTM 15.1 (the reference software for 3D-HEVC) [40]. Eight sequences recommended by JCT-3V in two resolutions (i.e., 1024×768 and 1920×1088) were used as inputs, and each sequence was composed of three views (three texture videos and associated depth maps). Tests were conducted at

both the common test condition (CTC) [41] and at the all-intra (AI) configuration, and the QP combinations for the texture video and the depth map were (25, 34), (30, 39), (35, 42), and (40, 45).

The performance results of the proposed fast mode decision algorithm and the original algorithm in HTM under the CTC configuration are compared in Table 5. "Texture 0", "Texture 1", and "Texture 2" represent the BD-rate [42] performance for the three texture videos. "Texture PSNR/texture bitrate" and "Synth PSNR/total bitrate" represent the BD-rate performance considering Y-PSNR of the coded texture views over the bit rate of texture data and Y-PSNR of the synthesized texture views over the bit rate of both texture and depth data, respectively. "ΔT" is the total reduction of the encoding time, and it was used to evaluate the computational complexity. It was derived by $\Delta T = (T_{org} - T_{prop})/T_{org} \times 100\%$, where T_{org} and T_{prop} are the execution time required for encoding the test sequences with the original 3D-HEVC encoder and with the encoder including the proposed fast mode decision algorithm, respectively. They are the geometric mean of four execution times for encoding the same sequences with different QP combinations. Table 5 shows that the average BD-rates for the three texture videos were 0, -0.07%, and -0.01%. "Texture PSNR/texture bitrate" and "Synth PSNR/total bitrate", the overall evaluation for texture videos and synthesized views, were -0.01% and 0.05%. These data indicate that the proposed algorithm had almost no effect on the coding performance of the texture videos and the synthesized views. However, it achieved a reduction in the total encoding time of 5.92%.

Table 5. Results of the proposed algorithm compared with the 3D-HEVC encoder under CTC

Sequence	Texture 0	Texture 1	Texture 2	Texture PSNR / texture bitrate(%)	Synth PSNR / total bitrate(%)	Δ T(%)
Balloons	0.00	-0.04	0.08	0.01	0.01	5.63
Kendo	0.00	-0.01	0.06	0.02	0.06	6.49
Newspaper_CC	0.00	0.03	0.09	0.03	0.01	4.35
GT_Fly	0.00	-0.18	-0.23	-0.03	0.04	6.20
Poznan_Hall2	0.00	-0.32	-0.03	-0.07	0.03	7.25
Poznan_Street	0.00	0.16	-0.08	0.00	0.05	4.85
Undo_Dancer	0.00	-0.04	0.02	-0.01	0.11	6.57
Shark	0.00	-0.17	0.05	-0.04	0.12	5.98
Average	0.00	-0.07	-0.01	-0.01	0.05	5.92

The comparative results of the AI configuration are shown in Table 6, and it can be see that the BD-rates for the texture videos (Texture 0, Texture 1, and Texture 2) were all 0 and that the "Synth PSNR/total bitrate" was 0.08%. At the same time, the average total encoding time was reduced by 31.12%. These results indicate that the proposed algorithm was more effective in terms of reducing the encoding time for the AI configuration and that it maintained almost the same coding performance as the original 3D-HEVC encoder. The AI configuration will be the focus of attention in the subsequent experiment.

In addition, the performance of our fast-mode decision algorithm was evaluated further by comparing it with the recent optimal pruning tree (OPT) method [37]. It is based on the SDI mode and aims to improve the intra-coding of the depth map. Table 7 shows that the OPT method achieved an average of 25.63% reduction in the total encoding time with a 0.18% BD-rate loss compared with the original 3D-HEVC encoder. By comparing

Table 6. Results of the proposed algorithm compared with the 3D-HEVC encoder under AI

Sequence	BD-rate(Texture 0, Texture 1, Texture 2)	Synth PSNR / total bitrate(%)	Δ T(%)
Balloons	0.00	0.03	27.29
Kendo	0.00	0.04	31.05
Newspaper_CC	0.00	0.02	20.01
GT_Fly	0.00	0.09	35.59
Poznan_Hall2	0.00	0.14	40.56
Poznan_Street	0.00	0.05	31.40
Undo_Dancer	0.00	0.13	31.02
Shark	0.00	0.12	32.01
Average	0.00	0.08	31.12

Tables 6 and 7, it is found that the proposed algorithm reduced further the total encoding time by 5.49%.

Table 7. Results of the OPT method compared with the 3D-HEVC encoder under AI

Sequence	BD-rate(Texture 0, Texture 1, Texture 2)	Synth PSNR / total bitrate(%)	Δ T(%)
Balloons	0.00	0.05	20.00
Kendo	0.00	0.12	22.63
Newspaper_CC	0.00	0.10	11.11
GT_Fly	0.00	0.10	20.74
Poznan_Hall2	0.00	0.40	45.34
Poznan_Street	0.00	0.28	17.87
Undo_Dancer	0.00	0.27	36.11
Shark	0.00	0.13	31.27
Average	0.00	0.18	25.63

As mentioned previously, the overall algorithm incorporates three key points. Then, the contribution of each key point in terms of encoding time reduction with different QP combinations was analyzed further. The first point is referred to as an aggregated RD cost method based on partial sub-CUs, and its contribution with respect to the reduction in the encoding time required is detailed in Fig 7, which shows that the reduction of the encoding time increased gradually with the QP. This indicates that this method will skip more sub-CU nodes and be more effective in the case of a large QP value.

The other two key points are based on the DIS mode, and they are referred to collectively as the DIS-based method. The curve of the reduction of encoding time benefited from this method is shown in Fig 8, which shows that the reduction in the encoding time decreased as QP increased. This is related to the working mechanism of the DIS mode. It is well known that the larger QP becomes, the worse the quality of the reconstructed signal will be. In the DIS mode, the current CU was predicted by using the neighboring reconstructed samples. Therefore, the DIS mode will no longer have the advantage in prediction accuracy over the other prediction modes in the case of large QP values. That is, the probability that the DIS mode eventually will be selected as the optimal prediction

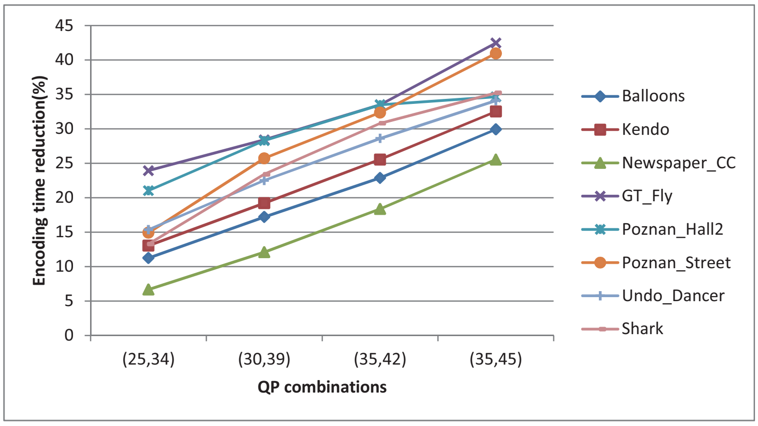

Figure 7. Curves of the reduction in encoding time based on the method of the Aggregated RD cost of partial

mode decreases as QP increases. As a result, the effectiveness of the proposed DIS-based method also will be reduced as the QP increases.

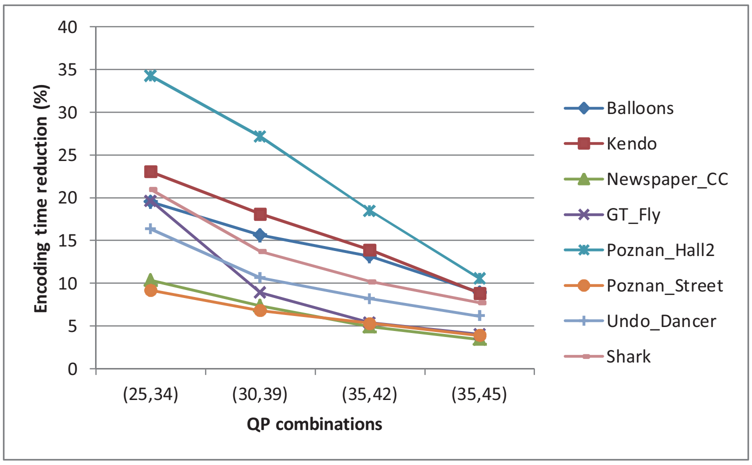

Figure 8. Curves of the reduction in encoding time provided by the DIS-based method

6. **Conclusions.** In this paper, we proposed a fast algorithm to reduce the computational complexity of the depth map mode decision in 3D-HEVC. The mode decision is performed by traversing all nodes in a complete quadtree in depth-first order, and every node gets an optimal PU mode by testing all PU modes. The proposed algorithm reduces the computational complexity by skipping the testing of part of the PU modes for a node, sometimes even skipping the whole node. The performance of the proposed algorithm was tested over a set of video sequences with different QP combinations, and it was compared with the original 3D-HEVC encoder and a similar method (OPT). The experimental results indicated that the proposed algorithm effectively reduced the coding complexity of the depth map, particularly under the AI configuration, while maintaining almost the same coding performance as the original 3D-HEVC encoder. In addition, it achieved a greater reduction in the complexity of the coding operation than the OPT method.

Acknowledgment. We would like to acknowledge the supports from the National Natural Science Foundation of China (No. 61210006 and 61672373), International Cooperative

Program of Shanxi Province (No. 2015081015), and Scientific and Technological project of Shanxi Province (2015031003-2).

REFERENCES

[1] K. Mller, H. Schwarz, Marpe D, Bartnik C, Bosse S, Brust H, Hinz T, Lakshman H, Merkle P, Rhee FH, Tech G, Winken M, Wiegand T, 3D High-Efficiency Video Coding for Multi-View Video and Depth Data. *IEEE Transactions on Image Processing* vol.22, no.9 pp.3366-3378

[2] C. Fehn, A 3D-TV approach using depth-image-based rendering (DIBR). in *Visualization, Imaging, and Image Processing*, pp.482-487

[3] P. Merkle, C. Bartnik, K. Muller, Marpe D, Wiegand T (2012) 3D video: Depth coding based on inter-component prediction of block partitions. in *Picture Coding Symposium (PCS)*, pp.149-152

[4] J.Y. Lee, M. W. Park, C. Kim C, 3D-CE1: Depth intra skip (DIS) mode. ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-K0033

[5] F. Jger, Simplified depth map intra coding with an optional depth lookup table. in *International Conference on 3D Imaging (IC3D)*, pp.1-4

[6] R. H. Gweon,Y. -L. Lee , Lim J, Early Termination of CU Encoding to Reduce HEVC Complexity. ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-F045

[7] J. Yang, J. Kim, K. Won, H. Lee, B. Jeon, Early SKIP Detection for HEVC. ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-G543

[8] K. Choi, S. H. Park , E. S. Jang, Coding tree pruning based CU early termination. ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-F092

[9] Q. Yu, X. Zhang, S. Wang, S. Ma, Early termination of coding unit splitting for HEVC. in *Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp.1-4

[10] J. Kim, Y. Choe, Y. G. Kim, Fast Coding Unit size decision algorithm for intra coding in HEVC. in *IEEE International Conference on Consumer Electronics (ICCE)*, pp.637-638

[11] Y. C. Lin, J. C. Lai, Edge Density Early Termination Algorithm for HEVC Coding Tree Block. in *Computer, Consumer and Control (IS3C), 2014 International Symposium on*, pp.39-42

[12] N. Shan, W. Zhou, Z. Duan, H. Wei, A Fast Coding Unit Depth Decision Algorithm for HEVC Inter Prediction. in *2015 Ninth International Conference on Frontier of Computer Science and Technology*, pp.316-320

[13] C. Zhou, F. Zhou, Y. Chen, Spatio-temporal correlation-based fast coding unit depth decision for high efficiency video coding. *ELECTIM*, vol.22, no.4 pp.043001-043001

[14] Y. Zhang, H. Wang, Z. Li, Fast Coding Unit Depth Decision Algorithm for Interframe Coding in HEVC. in *Data Compression Conference (DCC)*, 2013, pp.53-62

[15] L. Shen, Z. Liu , X. Zhang , W. Zhao, Z. Zhang , An Effective CU Size Decision Method for HEVC Encoders. *IEEE Transactions on Multimedia vol.15*, no.2 pp.465-470

[16] L. Shen, Zhang Z, Liu Z, Adaptive Inter-Mode Decision for HEVC Jointly Utilizing Inter-Level and Spatiotemporal Correlations. *IEEE Transactions on Circuits and Systems for Video Technology*, vol.24, no.10 pp.1709-1722

[17] Z. Pan, S. Kwong, Sun MT, Lei J, Early MERGE Mode Decision Based on Motion Estimation and Hierarchical Depth Correlation for HEVC. *IEEE Transactions on Broadcasting*, vol.60, no.2 pp.405-412

[18] J. Liu, Jia H, Xiang G, Huang X, Cai B, Zhu C, Xie D, An adaptive inter CU depth decision algorithm for HEVC. in *2015 Visual Communications and Image Processing (VCIP)*, pp 1-4

[19] G. Correa, Assuncao P, Agostini L, Cruz LAdS, Complexity control of HEVC through quadtree depth estimation. in *EUROCON, 2013 IEEE*, pp 81-86

[20] X. Shen, Yu L, Chen J, Fast coding unit size selection for HEVC based on Bayesian decision rule. in *Picture Coding Symposium (PCS)*, 2012, pp 453-456

[21] Cho S, Kim M, Fast CU Splitting and Pruning for Suboptimal CU Partitioning in HEVC Intra Coding. *IEEE Transactions on Circuits and Systems for Video Technology*, vol.23, no.9 pp.1555-1564

[22] K. Lim, Lee J, Kim S, Lee S, Fast PU Skip and Split Termination Algorithm for HEVC Intra Prediction. *IEEE Transactions on Circuits and Systems for Video Technology*, vol.25, no.8 pp.1335-1346

[23] X. Shen, Yu L, CU splitting early termination based on weighted SVM. *EURASIP Journal on Image and Video Processing*, 2013 (1 pp.1-11

[24] Y. C. Liu, Chen ZY, Fang JT, Chang PC, SVM-Based Fast Intra CU Depth Decision for HEVC. in *2015 Data Compression Conference*, pp.458-458
[25] T. Zhang, Sun MT, Zhao D, Gao W, Fast Intra Mode and CU Size Decision for HEVC. *IEEE Transactions on Circuits and Systems for Video Technology*, PP (99):1-1
[26] J. Xiong, Li H, Wu Q, Meng F, A Fast HEVC Inter CU Selection Method Based on Pyramid Motion Divergence. *IEEE Transactions on Multimedia*, vol.16, no.2 pp.559-564
[27] T. Zhao T, Wang Z, Kwong S, Flexible Mode Selection and Complexity Allocation in High Efficiency Video Coding. *IEEE Journal of Selected Topics in Signal Processing*, vol.7, no.6 pp.1135-1144
[28] J. Jung, Viswanathan K, Depth Quadtree Prediction for HTM. ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-A0044
[29] L. Shen, Zhang Z, Efficient depth coding in 3D video to minimize coding bitrate and complexity. *Multimedia Tools and Applications*, vol.72, no.2 pp.1639-1652
[30] Q. Zhang, Li N, Wu Q, Fast Mode Decision for 3D-HEVC Depth Intracoding. *The Scientific World Journal*, 2014:9
[31] L. Shen, An P, Zhang Z, Hu Q, Chen Z, A 3D-HEVC Fast Mode Decision Algorithm for Real-Time Applications. *ACM Trans Multimedia Comput Commun Appl*, vol.11, no.3 pp.1-23
[32] G. Zhouye, Jianhua Z, Nam L, Zhang P, Fast Depth Modeling Mode selection for 3D HEVC depth intra coding. in *Multimedia and Expo Workshops (ICMEW), 2013 IEEE International Conference on*, pp.1-4
[33] Z. Gu, Zheng J, Ling N, Zhang P, Fast Intra Prediction Mode Selection for Intra Depth Map Coding. ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-E0238
[34] G. Sanchez, Saldanha M, Balota G, Zatt B, Porto M, Agostini L, DMMFast: a complexity reduction scheme for three-dimensional high-efficiency video coding intraframe depth map coding. *ELECTIM*, vol.24, no.2 pp.023011-023011
[35] C. S. Park, Edge-Based Intramode Selection for Depth-Map Coding in 3D-HEVC. *IEEE Transactions on Image Processing*, vol.24, no.1 pp.155-162
[36] C. S. Park, Efficient intra-mode decision algorithm skipping unnecessary depth-modelling modes in 3D-HEVC. *Electronics Letters*, vol.51, no.10 pp.756-758
[37] K. Miok, Nam L, Li S, Fast single depth intra mode decision for depth map coding in 3D-HEVC. in *Multimedia & Expo Workshops (ICMEW), 2015 IEEE International Conference on*, pp.1-6
[38] P. Merkle, Rhee H, Mueller K, Wiegand T, Fast Wedgelet search. ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-A0105
[39] H. Zhang, Ma Z, Fast Intra Mode Decision for High Efficiency Video Coding (HEVC). *IEEE Transactions on Circuits and Systems for Video Technology*, vol.24, no.4 pp.660-668
[40] 3D-HEVC reference software: 'Joint collaborative team on 3D video (JCT-3 V), HTM-15.1 (2015). https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-15.1/
[41] D. Rusanovskyy, Mller K, Vetro A, Common Test Conditions of 3DV Core Experiments. ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, JCT3V-F1100
[42] G. Bjontegaard, Calculation of average PSNR differences between RD-curves. ITU-T SG16 Q6, VCEG-M33