IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 29, 2020, accepted July 12, 2020, date of publication July 15, 2020, date of current version July 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3009424

A Fast and Efficient 3D-HEVC Method for Complexity Reduction Based on the Correlations of Inter-View, Spatio-Temporal, and Texture-Depth

QIUWEN ZHANG[1], (Member, IEEE), YIHAN WANG[1], TAO WEI[2], LIXUN HUANG[1], AND RIJIAN SU[1]
[1]College of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China
[2]Computer College, Henan Institute of Engineering, Zhengzhou 450001, China

Corresponding author: Rijian Su (rijiansu@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61771432, Grant 61773018, Grant 61302118, Grant 61702464, Grant 61374014, and Grant 61803346.

ABSTRACT 3D High Efficiency Video Coding (3D-HEVC) is the latest High Efficiency Video Coding (HEVC) standard development to improve the compression performance of multi-view video plus depth (MVD) format. It adopts several additional coding tools for better representation of the dependent texture and depth data. These tools achieve high compression efficiency, but require high computational complexity. In this paper, we introduce a fast 3D-HEVC method for the complexity reduction of MVD. A simulation analysis is performed to study the coding information correlations of inter-view, spatio-temporal and texture-depth. The proposed fast algorithm decides to skip some treeblocks of texture and depth map at the early stage without a normal coding process, which includes fast depth level, early SKIP/Merge mode, adaptive early termination, based on the correlations of inter-view, spatio-temporal, and texture-depth. Experimental results show that the proposed fast scheme can save 70.2% encoding runtime of 3D-HEVC, with negligible loss of rate-distortion (RD) performance.

INDEX TERMS 3D-HEVC, low complexity, texture video, depth map.

I. INTRODUCTION

3D video is getting more and more popular due to the recent advances in stereoscopic display technologies and related applications, such as 3D television (3DTV), 3D film and 3D game. The next generation 3D representation, MVD [1], which consists of texture image and associated depth map, has been recommended by Moving Picture Experts Group (MPEG). It enables receiver to synthesized intermediate views by using depth image based rendering (DIBR) [2]. To improve the coding efficiency of MVD, the 3D-HEVC has been developed by the latest video standard-HEVC [3]–[5]. Research effort in both industry and academia have been put together to better understand and explore the texture videos, corresponding depth data characteristics and developing efficient MVD coding methods.

To efficient compression of MVD data, the additional tools are designed to exploit the dependencies between the components. The official 3D-HEVC reference software can achieve over 46% bit-rate savings over HEVC simulcast for 3D video contents [6]. This significant compression improvement is attained from several coding tools beyond conventional single-layer video coders HEVC, known as "Neighboring block disparity vector (NBDV)", "Inter-view motion prediction (IMP)" and "View synthesis prediction (VSP)" for texture videos, "Depth modeling modes(DMMs)", "Segment-wise DC coding (SDC)" and "Motion parameter inheritance (MPI)" for depth maps, respectively. However, these additional coding tools bring significant computation complexity of the encoder, and actually handicap the chances of application. The complexity

The associate editor coordinating the review of this manuscript and approving it for publication was Chao Tan.

of 3D-HEVC becomes a critical importance issue. Therefore, it needs a further reduction on computational complexity with negligible loss of coding performance to promote 3D-HEVC broader applications.

In order to simplify the previous 3D video coding standard with large complexity, some fast methods [7]–[13] are proposed. A view-adaptive algorithm for fast motion estimation (ME) based on the motion information and the complexity of candidate modes is investigated in [7], which can achieve the computational complexity reduction of multi-video coding (MVC). An early decision method is introduced in [8] to balance between time reduction and coding performance of MVC, which considers both probability and coding time of all coding modes, a novel model is employed by taking selection method into considering based on the multi-view coding mode characteristics. An early decision method utilizing the advantage of inter-component correlation is proposed in [9], which can achieve time saving for coding process. In [10], a fast decision is made to decrease candidate modes by the relationship between the current and adjacent view. The fast macroblock encoding method in [11] reduces the number of candidate modes: fast mode decision. Firstly, the procedure of SKIP/Direct mode detection can extract RD information of the current block. Secondly, the ME/disparity estimation (DE) and mode decision process are optimized to utilize the obtained RD activity. Finally, the spatial and inter-view correlations are applied to perfect the coding performance. A fast decision algorithm is developed in [12] to accelerate the encoding process of MVC based on high and directional correlation among variable block size coding modes. Adaptive DE and fast mode size decision are developed in our former work [13] by utilizing the correlation of texture image and depth map to achieve the complexity reduction of MVD. However, these fast methods are not suitable for the HEVC-based 3D video encoder, which employs new quadtree structured coding unit (CU) and additional coding tools with increased complexity in 3D-HEVC. Furthermore, the previous fast MVC methods are more favorably designed for the multi-view video compression with lower spatial resolution.

Recently, some fast algorithms of HEVC also have been developed in the literatures. They consist of two different works: methods with intra mode decision and CU size prediction. Category 1: fast intra mode decision methods are designed in literatures [14]–[17]. A fast decision method is investigated in [14] based on the direction information among the current CU and neighboring CUs. In [15], fast intra mode decision method adopts the edge information contained in prediction unit (PU) to achieve candidate prediction directions reduction. In [16], a complexity control method can dynamically adjust the partitioning level of the CU that is defined by quadtree-based structure through utilizing fast mode decision algorithm. An early CU size decision is designed in [17] for fast intra prediction. The aforementioned fast intra mode decision methods are well researched for HEVC to reduce coding time with acceptable quality degradation. However, since the coding structure involved in disparity compensated and mode prediction is different from that of HEVC, it is necessary to further improve the performance of mode determination processing in 3D-HEVC. Category 2: methods with fast CU size prediction are developed in literatures [17]–[25]. A fast partition method is designed in [18], which can achieve the complexity reduction of HEVC. A fast CU size decision is presented in [19] to reduce the candidate depth levels for intra prediction process of HEVC. A fast CU depth decision is developed in [20] to speed up the encoding time by utilizing the correlation of depth information within the spatio-temporal treeblock and the current treeblock. In [21], a mode selection method based on linear programming is reported to save computational complexity by the frames. A novel zero block (ZB) detection method is introduced in [22] to detect pseudo ZBs, and then avoid the increasing of coding time for original encoder. In [23], a CU selection algorithm can early skip the specific inter CUs by utilizing the pyramid motion divergence in HEVC. A fast mode [24] and CU size decision [25] will accelerate the coding process by jointly utilizing correlation between the spatio-temporal and quadtree structure, and the coding information introduced in referenced coded CUs. In [26], a fast transform unit (TU) size decision method is introduced to select the candidate transform sizes by using the Bayesian theory. A fast decision based on the correlation of spatio-temporal is investigated in our former work [27] as well, which achieves runtime reduction of HEVC. All these fast CU size prediction methods can effectively reduce complexity with small performance loss. However, these fast mode decision and CU size prediction methods are only designed for the texture video, which do not exploit the depth map properties or the characteristics of the new intra prediction modes (such DMMs). They are further used to reduce the complexity of texture and depth coding.

To address the limitation, some state-of-the-art works [28]–[39] on fast encoding have been presented for 3D-HEVC. These methods can be categorized into the following categories: Category 1: reducing the complexity of depth coding. A fast decision method is designed in [28] to allocate DMM complexities by classifying edge into different strategies, which can adaptively skip useless DMMs mode. A fast intra coding is introduced in [29] to early terminate the quadtree partition of depth map, which can adaptively detect corner point and reallocate the partitioning level. A fast mode decision is presented in [30] to reduce the candidate modes in depth coding process, where the correlation between inter-view and grayscale similarity is used to search the optimal PU mode. A fast decision [31] can save the encoding time of depth map by using the correlation of depth-texture and edge information. It can omit several characteristics in depth levels and prediction modes, which rarely appear in mode decision. A fast 3D-HEVC depth map compression is exploited in [32] to correlate the rate-distortion optimization (RDO) process, which is based on machine learning and data mining to extract correlations among the 3D-HEVC context attributes. It reports averaged 58.3% complexity reduction with 0.13% bit-rate loss for 3D-HEVC. A fast depth map intra mode decision method is developed in [33] to assign computational complexity based on tensor feature extraction and data analysis, which can selectively omit unnecessary depth modes in the 3D-HEVC mode decision process.

Category 2: reducing the complexity of texture video coding. An efficient encoder is presented in [34] and [35] to speed up the texture video compression, two strategies are employed by utilizing the correlation of inter-view. It reports averaged 47.1% time saving with 0.1% bit-rate loss. A fast texture video mode decision method based on the depth information is investigated in [36], which can achieve the computational complexity reduction of 3D-HEVC. This method exploits the correlations of depth map values to simplify the coding process of texture video. In [37], an online-learning-based method which is designed can speed up the texture coding, where this algorithm can adjust the prediction modes search in texture video coding.

Category 3: reducing the complexity of both texture and depth coding. A novel mode decision method is proposed in [38] to accelerate the most time-consuming prediction process, which adaptively adjusts the mode decision procedure by using the correlation of adjacent CUs depth levels and the correlation of texture-depth. Up to 66.0% peak complexity reduction is reported over HTM 9.0 with about 1.3% sacrifice for the synthesized video quality. An efficient CU partition process method is employed in [39] to save coding time for real-time applications. A fast mode decision method is introduced in [40] to save the coding time of 3D-HEVC by using the gradient information. In our previous work [41], a fast MVD coding scheme including fast motion and mode size selection is presented for 3D-HEVC, where about 60.1% complexity reduction is achieved with only 0.57% BD-rate loss. The aforementioned methods can effectively accelerate the encoding time of 3D-HEVC and maintain almost the same video quality with original encoder. However, most of the existing methods only optimize the correlations among inter-view, inter-component and the inter-level, while the coding information among mode prediction size, RD cost and CU partition characteristics in 3D-HEVC is not fully explored. Although a few fast coding algorithms are proposed to select proper prediction directions for previous compression standard as H.264, HEVC and MVC, there is still much room to further reduce the coding complexity of 3D-HEVC.

In this paper, a novel fast algorithm is presented to decrease the computational complexity of 3D-HEVC. View similarities exist in the texture and depth image contents (instant and view point at the same time), which contains strong correlation. In our previous works [41] and [42], we skipped some unnecessary modes in 3D-HEVC to reduce the coding complexity only according to the correlations of temporal-spatial and inter-view. As a result, we cannot achieve the optimum coding performance, which has a relatively low the correlation of texture-depth. Hence, the proposed algorithm reduces the coding complexity by skipping unnecessary modes under the consideration of texture-depth as well as temporal-spatial and inter-view correlations. The main idea of this paper is jointly exploiting coding information of inter-view, spatio-temporal, and texture-depth to extract mode prediction features for 3D-HEVC, and then optimizing prediction modes of the current texture and depth treeblocks. It comprises three approaches: fast CU depth level range decision method, early SKIP/Merge mode detection method, and adaptive early termination mode prediction method. The main contributions of the proposed method are as follows: (1) unlike the formerly fast CU depth decision, fast CU depth level range determination makes decision based on a motion complexity parameter, which is defined according to the correlations of motion information not only among spatial-temporal neighboring treeblocks, but also between previously coded view and texture-depth; (2) early SKIP and Merge mode detection reduces unnecessary prediction modes according to the mode context of inter-view, spatio-temporal and texture and depth video neighboring treeblocks; (3) adaptive early termination mode prediction uses the RD cost correlations between the coding information of inter-view, spatio-temporal, texture-depth neighboring CUs, and the previously coded treeblocks to exploit the adaptive threshold for fast coding. As far as we know, there is almost no work similar to the proposed algorithm on the 3D-HEVC.

The rest of the article is organized as follows. In section II, the observations and analysis are introduced. Section III explains the proposed fast method in detail. Section VI provides the performance of the proposed method. Finally, Section V concludes this paper.

II. OBSERVATIONS AND ANALYSIS

Since the additional coding tools are introduced into 3D-HEVC, specifically used to depth map, which has the same quadtree structure with HEVC. The mode prediction process of 3D-HEVC is performed by using all coding modes to select the best one with minimum RD cost. Therefore, the cost function J_{mode} is evaluated as,

$$J_{\text{mode}} = (SSE_{luma} + \omega_{chroma} \cdot SSE_{chroma}) + \lambda_{\text{mode}} \cdot R_{\text{mode}} \quad (1)$$

where SSE_{luma} and SSE_{chroma} are the distortion between current treeblock and reconstructed treeblock on luma and chroma component, respectively. ω_{chroma} is the weighing parameter, λ_{mode} is the Lagrange multiplier, R_{mode} is total bit-rate cost. This "try all then select the best" method can improve coding efficiency of encoder, but requires a large complexity. Furthermore, the additional new prediction tools (such as DMMs, SDC) have been developed for depth image to improve the coding efficiency of MVD. Compared with HEVC, all the new tools are added to the full RD search list in 3D-HEVC. These techniques will make high complexity and obstruct the application of 3D-HEVC.

In video sequences, it exists high correlation between the adjacent treeblocks of spatial and temporal. The coding message of the current treeblock is highly inseparable to

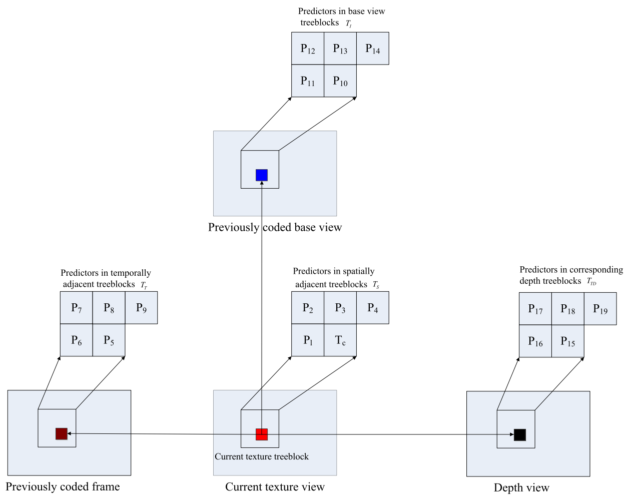


FIGURE 1. **Predictors of current CU.**

the prediction mode of its spatial and temporal neighboring treeblocks. Therefore, we can exploit the correlations of spatial and temporal to analyze the current treeblock features and skip certain needless modes. Furthermore, the correlation of inter-view is developed to reduce data redundancy of MVD in HTM encoders. View similarities exist in the texture video and depth map, which leads to strong relation among these multi-view contents (the relationship between the current treeblock and the corresponding treeblocks as shown in Fig. 1). After the above observations, the aforementioned correlations among the inter-view are analyzed on some video test sequences. Among the video sequences, “*Kendo*”, “*Balloons*” and “*Newspaper*” are 1024 × 768 resolution, “*Undo_Dancer*”, “*GT_Fly*”, “*Poznan_Street*”, “*Poznan_Hall2*” and “*Shark*” are 1920 × 1088 resolution. Test conditions are set as follows: 3-view case, quantization parameters (QPs) of both texture and depth are set as 25, 30, 35 and 40, VSP and DMMs are on.

TABLE 1. **Coding information correlation.**

Sequences	Texture video		Depth map	
(%)	Same CU depth level	Other	Same CU depth level	Other
Kendo	76.1	19.9	73.3	26.7
Balloons	79.2	20.4	72.4	27.6
Newspaper	74.8	22.2	70.8	29.2
Shark	71.2	27.5	66.9	33.1
Undo_Dancer	72.3	27.4	67.3	32.7
GT_Fly	78.4	21.6	71.8	28.2
Poznan_Street	80.2	19.8	73.4	26.6
Poznan_Hall2	82.7	16.3	80.6	19.4
Average	78.1	21.9	72.1	27.9

Table 1 shows the coding information correlations between current treeblock and its corresponding treeblock in adjacent independent views, in which “same CU depth level” denotes that current treeblock has same depth level as the neighboring views (suitable for both texture video and depth map). Table 1 demonstrates that the “same CU depth level” has more probability to be selected compared with the other different depth levels, and the average probability is 78.1% for overall texture video treeblocks, and 72.1% for overall depth map treeblocks, respectively. Therefore, CU depth level in current treeblock is decided based on the corresponding treeblock in adjacent independent view.

Meanwhile, the strong coding information correlations are obvious in multi-view video and depth. Since the texture and depth image in MVD represent the same scene, the coding characteristics may be similar in mode decision process. Table 2 shows the coding correlations between current treeblock of depth map and its co-located treeblock in texture video, where “TDL0”, “TDL1”, “TDL2” and “TDL3” are the depth levels of co-located treeblock of texture video, “same CU depth level” denotes that the current depth treeblocks have the same depth value as the co-located treeblock of texture. As shown in Table 2, when the texture video treeblocks choose the depth value is “0” (TDL0), the average probability of depth map treeblock which choose the depth value “0” (same CU depth level) reaches 96.5%; when the depth value of texture video treeblock is “1” (TDL1), “2” (TDL2) or “3” (TDL3), the probability of depth map treeblock with same CU depth level are more than 36.6%. This results indicate that the optimal depth value of texture video treeblock has a high possibility to be the best depth value of depth map treeblock. Thus, it takes advantage of coding information similarities between depth and texture to save the encoding time of 3D-HEVC.

Based on above observations, the coding correlations from the spatial-temporal, formerly coded view and texture-depth are utilized in this section. A set of coding prediction variables (ψ) is defined as,

$$\psi = \{T_S, T_T, T_I, T_{TD}\} \tag{2}$$

where T_S represents spatial predictors (including P_1, P_2, P_3 and P_4 in Fig.1), T_T represents temporal predictors (including P_5, P_6, P_7, P_8 and P_9) located at the same position in the current texture treeblock T_C, T_I represents the inter-view predictors in (including $P_{10}, P_{11}, P_{12}, P_{13}$ and P_{14}) in the neighbor coded view, T_{TD} represents the texture-depth predictors (including $P_{15}, P_{16}, P_{17}, P_{18}$ and P_{19}) from corresponding depth map view in Fig.1. Based on coding information in the predictors ψ, the current treeblock is extracted to skip unnecessary variable-size prediction modes.

III. PROPOSED FAST 3D-HEVC ALGORITHM FOR MVD

A. FAST CU DEPTH LEVEL RANGE DETERMINATION

In 3D-HEVC, the quadtree structure is also used to compress texture and depth data. In addition, CU depth level has a fixed range for all the texture and depth coding. Similar to HEVC, both ME and DE process will search all the candidate depth levels and find the best mode with minimum RD cost. The above techniques improve coding efficiency as high as

TABLE 2. **Depth level correlation of the depth map and texture video.**

Sequences	TDL0		TDL1		TDL 2		TDL3	
(%)	Same CU depth level	Other	Same CU depth level	Other	Same CU depth level	Other	Same CU depth level	Other
Kendo	96.8	3.2	61.2	38.8	45.8	54.2	34.2	65.8
Balloons	94.3	5.7	59.8	40.2	43.2	56.8	36.8	63.2
Newspaper	95.7	4.3	60.3	39.7	42.6	57.4	37.4	62.6
Shark	98.2	1.8	63.1	36.9	48.7	51.3	40.3	59.7
Undo_Dancer	97.8	2.2	62.9	37.1	48.3	51.7	40.7	59.3
GT_Fly	93.9	6.1	58.6	41.4	41.8	58.2	38.2	61.8
Poznan_Street	96.1	3.9	60.8	39.2	46.2	53.8	33.8	66.2
Poznan_Hall2	99.2	0.8	55.9	44.1	39.7	60.3	31.3	68.7
Average	96.5	3.5	60.3	39.7	44.5	55.5	36.6	63.4

possible in HTM, but bring significant computational complexity. In fact, small depth values occur very frequently for treeblocks in static region. On the other hand, small depth values are rarely selected for treeblocks with complex motion region [25]. The exhaustive mode decision is unnecessary, CU depth level is adaptively skipped by utilizing the motion characteristics of texture and depth treeblock.

Based on the motion predictors ψ of Equation (2), we use the motion information from the referenced predictors to explore ME characteristics of the current treeblock. We present a new standard to recognize motion complexity of among the current treeblock and the neighboring treeblocks. Motion vectors from the current texture treeblock and these from 4 × 4 block level covered by corresponding treeblocks $(P_1, P_2, \ldots, P_{19})$ are used to evaluate motion homogeneity. If motion vectors of the current texture treeblock and covered 4×4 block of corresponding treeblocks are defined as $MV_{ij} = (MVx_{ij}, MVy_{ij})$, the motion complexity in the horizontal x and vertical y directions are defined as,

$$MCx = \frac{1}{T}\sum_{(i,j)\in\psi} \left|MVx_{ij} - \rho_{ij}\cdot \sum_{(i,j)\in\psi} MVx_{ij}\right| \tag{3}$$

$$MCy = \frac{1}{T}\sum_{(i,j)\in\psi} \left|MVy_{ij} - \rho_{ij}\cdot \sum_{(i,j)\in\psi} MVy_{ij}\right| \tag{4}$$

where MCx and MCy denote the motion complexity of the horizontal x and vertical y, respectively. T is the total amount (4 × 4 block) of current treeblock and its adjacent treeblocks, ρ_{ij} denotes weight factor, which depends on their correlation from the current CU. The stronger correlation among the current CU and adjacent treeblocks, the larger of the value should be assigned. According to the extensive experiments, the weights for adjacent treeblocks in the spatial predictors are set to 0.4, and the weight for temporal, inter-view and texture-depth predictors are set to 0.2. In this paper, the motion complexity parameter is defined in function (5),

$$MC = MCx + MCy \tag{5}$$

Based on the motion complexity parameter MC, the current treeblock T_c is classified into two types: static and complex motion treeblocks, which is defined as follows,

$$T_c \subset \begin{cases} Static\ region & if\ MC < R \\ Complex\ motion\ region & otherwise \end{cases} \tag{6}$$

where R denotes the threshold factor, which determines the treeblock with static or complex motion region. The value of R is crucial for mode decision in 3D-HEVC, and it can balance complexity reduction and coding quality. Based on simulations results, it can be found that the optimal value of R depends on each sequence content. The threshold R is calculated by analyzing the motion homogeneity of a treeblock in each sequence. R is calculated by motion vectors, so as long as motion vectors are obtained, which is calculated without coding. Assume that a treeblock located at the a th row and b th column, the R is defined for each sequence as:

$$R = 0.8 \times \frac{1}{\mathrm{A}\times\mathrm{B}} \cdot \sum_{a=0}^{\mathrm{A}-1}\sum_{b=0}^{\mathrm{B}-1}\left(MVx_{a,b} + MVy_{a,b}\right) \tag{7}$$

where A and B represent the number of treeblocks of a row and a column in each test frame, respectively.

TABLE 3. **CU depth level distribution for different treeblocks in texture coding.**

Sequences	Treeblocks in static region				Treeblocks in complex motion region			
Level(%)	0	1	2	3	0	1	2	3
Kendo	68.7	29.6	1.1	0.6	1.2	5.3	26.9	66.6
Balloons	71.2	27.4	0.9	0.5	1.8	5.6	27.6	65.0
Newspaper	64.2	31.9	2.7	1.2	2.4	5.9	30.7	61.0
Shark	59.6	38.4	1.3	0.7	0.9	3.8	26.1	69.2
Undo_Dancer	60.8	36.7	1.6	0.9	0.8	3.6	25.8	69.8
GT_Fly	67.2	30.1	1.7	1.0	1.3	5.4	28.7	64.6
Poznan_Street	69.5	28.6	1.2	0.7	1.5	4.8	30.1	63.6
Poznan_Hall2	74.6	24.5	0.6	0.3	2.2	6.7	31.8	66.6
Average	67.0	30.9	1.4	0.7	1.5	5.1	28.5	65.8

We analyze the CU depth level distribution for two types of treeblocks in Tables 3 and 4. "Level 0", "Level 1", "Level 2" and "Level 3" are the CU depth levels of texture and depth treeblock. Tables 3 and 4 show that the CUs with static region, its average probabilities of choosing "Level 0"

TABLE 4. CU depth level distribution for different treeblocks in depth map coding.

Sequences	Treeblocks in static region				Treeblocks in complex motion region			
Level(%)	0	1	2	3	0	1	2	3
Kendo	94.1	4.3	1.2	0.4	4.6	18.9	32.8	43.7
Balloons	96.3	2.8	0.6	0.3	5.1	19.2	33.2	42.5
Newspaper	92.1	5.2	1.9	0.8	3.9	21.6	34.6	39.9
Shark	91.4	5.7	2.4	0.5	3.4	16.9	29.1	50.6
Undo_Dancer	91.6	5.5	2.2	0.7	3.7	17.2	31.3	47.8
GT_Fly	93.8	3.9	1.7	0.6	4.7	19.6	33.8	41.9
Poznan_Street	95.4	3.2	1.1	0.3	4.2	18.4	32.7	44.7
Poznan_Hall2	97.1	2.3	0.4	0.2	5.8	22.6	36.7	34.9
Average	94.0	4.1	1.4	0.5	4.4	19.3	33.0	43.3

TABLE 5. CU depth levels for texture and depth treeblocks.

Treeblock type	Texture video		Depth map	
	Candidate depth levels	CU depth range $[Depth_{\min}, Depth_{\max}]$	Candidate depth levels	CU depth range $[Depth_{\min}, Depth_{\max}]$
Static region	0,1	[0,1]	0	[0,0]
Complex motion region	2,3	[2,3]	1,2,3	[1,2,3]

and "Level 1" are 67.0%, and about 30.9% in texture video, and the average probability of remaining CUs depth level is 1.4%. Therefore, when the depth level is set to "0" and "1", it will most likely cover about 97.9% texture video treeblocks. The CU mode decision on depth level "2" and "3" is skipped. Correspondingly, about 94.0% of depth map treeblocks choose the optimal CU depth level with "0", and total of "Level 1", "Level 2" and "Level 3" is less than 6.0%. Thus, the CU mode prediction with depth level of "1", "2" and "3" is skipped. For treeblocks with complex motion region, the probabilities of choosing "Level 2" and "Level3" are 28.5% and 65.8% in texture video, respectively, and total of "Level 0" and "Level 1" is less than 6.6%. If the CU depth level is set to "2" and "3", it will be cover about 93.4% of texture treeblocks. Meantime, the probability of depth level "1" in depth map is not very low, more than 19.3%, and thus only CU mode decision on depth level of "0" is skipped in depth compression. Therefore, the above analysis shows that the CU depth levels that are performed in 3D-HEVC are summarized in Table 5.

B. EARLY SKIP AND MERGE MODE DETECTION

In 3D-HEVC, the prediction modes for treeblock in inter-frame include SKIP, Merge, inter 2N × 2N, 2N × N, N × 2N, N × N, 2N × nU, 2N × nD, nL × 2N, nR × 2N and Intra 2N × 2N, N × N. When a treeblock chooses SKIP and Merge mode, it has no significant residual coefficients. Thus, SKIP and Merge mode provide outstanding coding performance with a little complexity, once SKIP and Merge mode are early determined, the exhaustive ME and DE process are omitted in mode decision. But the aforementioned decision procedure is delayed since the computation of RD cost in original encoder needs to fully search and then choose the SKIP/Merge mode with least cost. Since the nature texture video possess substantial amounts of background and static regions and depth map takes advantage of nearly constant and homogeneous regions, numerous CUs will choose SKIP and Merge mode based on RD cost. Therefore, if Merge/SKIP mode is decided in early mode decision process, large coding time could be reduced in original encoder. Based on above analysis, we present an early SKIP and Merge mode detection to omit ME and DE process in 3D-HEVC.

In order to detect SKIP/Merge mode correlation between current treeblock and treeblocks in predictors ψ, we classify the current treeblock into two types, i.e., "$TB1$" and "$TB2$". "$TB1$" denotes the neighboring treeblocks (including T_S, T_T, T_I, T_{TD} in Fig.1), which chooses SKIP/Merge mode, and "$TB2$" represents the remaining treeblocks. "Whole" indicates the total number of treeblocks.

Tables 6 and 7 show the SKIP and Merge mode distribution for different types of treeblocks with full mode decision. As shown in Tables 6 and 7, SKIP and Merge mode with a higher possibility are selected, and the average probability reaches 63.8% for texture coding, and about 82.7% for depth coding in whole treeblocks. For type of "$TB1$" treeblock, about 96.6% of texture treeblock and about 98.1% of depth treeblock will choose SKIP and Merge mode. For type "$TB2$" treeblock, about 32.5% of texture treeblock and 34.9% of depth treeblock will choose SKIP and Merge mode. In addition, we can find that many treeblocks are type "$TB1$", the ratio of type "$TB1$" is 49.1% for texture coding, and 75.1% for depth coding. Thus, if the corresponding predictors in ψ are coding with SKIP and Merge mode, it can be considered that two modes have high possibility to select as the optimal mode. Furthermore, if we can early detect SKIP and Merge mode for type "$TB1$" treeblock, a lot of encoding time can be omitted.

C. ADAPTIVE EARLY TERMINATION MODE PREDICTION

In 3D-HEVC, the mode prediction procedure finds the best one from all candidate modes by the RDO. In fact, ME/DE process on small CU size and mode prediction would be unnecessary in most cases because the treeblocks have substantial amounts of background or homogeneous region [39]. Therefore, it is better to propose an early termination strategy in 3D-HEVC. As mentioned above, the coding information of inter-view, spatio-temporal and texture-depth are strongly correlated to the current treeblock in mode decision procedure of 3D-HEVC. The results show that RD cost of the current treeblock is intimately related to neighboring treeblocks. Based on above analysis, the RD cost of the current treeblock $(RD\cos tT_{predict})$ is determined by using RD cost

TABLE 6. SKIP and merge distribution for texture video treeblocks.

Sequences	$TB1$		$TB2$		Whole		
(%)	SKIP mode and Merge mode	Other mode	SKIP mode and Merge mode	Other mode	SKIP mode and Merge mode	Other mode	Ratio of $TB1$
Kendo	96.3	3.7	34.1	65.9	67.3	32.7	53.4
Balloons	97.4	2.6	33.6	66.4	66.1	33.9	50.9
Newspaper	95.8	4.2	31.2	68.8	62.6	37.4	48.6
Shark	95.2	4.8	25.7	74.3	52.9	47.1	39.2
Undo_Dancer	95.1	4.9	25.3	74.7	52.5	47.5	38.9
GT_Fly	96.8	3.2	33.9	66.1	66.9	33.1	52.4
Poznan_Street	97.3	2.7	36.3	63.7	69.1	30.9	53.7
Poznan_Hall2	98.9	1.1	40.2	59.8	73.1	26.9	56.1
Average	96.6	3.4	32.5	67.5	63.8	36.2	49.1

TABLE 7. SKIP and merge distribution for depth map treeblocks.

Sequences	$TB1$		$TB2$		Whole		
(%)	SKIP mode and Merge mode	Other mode	SKIP mode and Merge mode	Other mode	SKIP mode and Merge mode	Other mode	Ratio of $TB1$
Kendo	98.3	1.7	38.2	61.8	84.4	15.6	76.9
Balloons	98.6	1.4	37.6	62.4	83.5	16.5	75.3
Newspaper	97.4	2.6	33.4	66.6	81.0	19.0	74.3
Shark	96.9	3.1	25.3	74.7	75.5	24.5	70.1
Undo_Dancer	96.6	3.4	25.1	74.9	74.9	25.1	69.7
GT_Fly	98.4	1.6	38.2	61.8	83.8	16.2	75.8
Poznan_Street	98.7	1.3	39.1	60.9	85.3	14.7	77.6
Poznan_Hall2	99.6	0.4	42.3	57.7	88.9	15.6	81.3
Average	98.1	1.9	34.9	65.1	82.7	18.3	75.1

of neighboring predictors ψ, (8), as shown at the bottom of the page, where i is the number of neighboring treeblocks, $RD\cos ti$ is RD cost of neighboring treeblocks, α_i is weight parameter of treeblock, which is allocated to the adjacent treeblocks according to their correlation from current treeblock. ξ_i is an adjust factor. Hence, ξ_i is set to "1", when adjacent treeblocks is available; otherwise, ξ_i is "0". χ, δ, ε and γ are the mode-weight factor. Similar to the previous equations (3) and (4), the stronger correlation between the current CU and adjacent treeblocks, the larger of the mode-weight factor value should be assigned. In this experiment, χ, δ, ε and γ are set to 0.4, 0.2, 0.2 and 0.2, respectively.

The threshold (THr) for early termination mode decision is calculated based on the minimum RD cost among $RD\cos tT_{predict}$ and spatial adjacent treeblocks,

$$THr = \mu \cdot \{RD\cos tP_1, RD\cos tP_2, RD\cos tP_3, RD\cos tP_4, RD\cos tT_{predict}\} \quad (9)$$

where μ represents an adjust parameter. When the minimum RD cost of the current treeblock (texture video or depth map) is smaller than THr, some ME and DE process on unnecessary CU sizes are skipped.

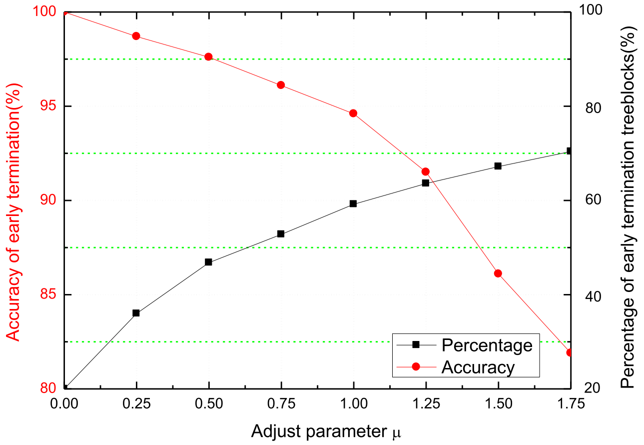


FIGURE 2. Relation in texture video coding.

The μ should be greatly save the encoding time of 3D-HEVC while it keeps a high accuracy. The accuracy of early termination (o) is evaluated as:

$$o = {}^{N_E}\!/_{N_F} \quad (10)$$

where N_E is the number of early terminated CUs, and N_F is the number of CUs which have the same optimal mode selected by full mode decision. Figs. 2 and 3 show the relationship between the value of μ and the percentage of

$$RD\cos tT_{predict} = \frac{\chi \cdot \sum_{i=1}^{4} \alpha_i \cdot \xi_i \cdot RD\cos ti + \delta \cdot \sum_{i=5}^{9} \alpha_i \cdot \xi_i \cdot RD\cos ti + \varepsilon \cdot \sum_{i=10}^{14} \alpha_i \cdot \xi_i \cdot RD\cos ti + \gamma \cdot \sum_{i=15}^{19} \alpha_i \cdot \xi_i \cdot RD\cos ti}{\sum_{i=1}^{19} \alpha_i \cdot \xi_i} \quad (8)$$

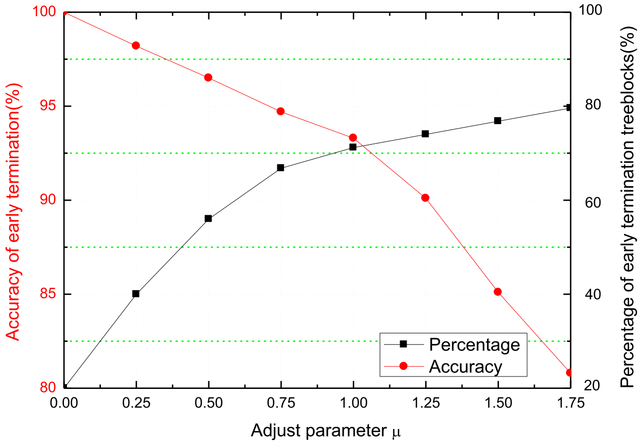


FIGURE 3. Relation in depth map coding.

early terminated CUs/accuracy of early termination in texture and depth, respectively. Eight sequences with different content characteristics are tested, including "Kendo", "Balloons", "Newspaper", "Undo_Dancer", "GT_Fly", "Poznan_Street", "Poznan_Hall2" and "Shark".

It is informed from Fig. 2 that for a treeblock of texture video, when $\mu = 1.30$, about 63% treeblocks of texture image can be early terminated, and the average accuracy of early determination is larger than 91%. When $\mu > 1.30$, the percentage of early terminated treeblocks of texture video slowly increases. At the same time, the early termination accuracy will decrease with the increase of μ, which leads to significant decrease in coding efficiency. When $\mu < 1.30$, the percentage of early terminated treeblocks of texture video and accuracy of early termination increase quite slowly. Thus, the threshold μ is set to 1.30 for texture coding in this paper. For a treeblock of depth map in Fig. 3, when $\mu = 1.30$, more than 76% treeblocks of texture video can be early terminated, and the accuracy of the early determination is larger than 93%. Thus, the value of the threshold μ for depth map coding is set to 1.05 in our work.

D. PROCEDURE OF THE PROPOSED OVERALL METHOD

Based on strategies of fast CU depth level, early SKIP and Merge mode detection and adaptive early termination mode prediction, the proposed fast 3D-HEVC method is described as follows:

Step 1) Start mode decision for a treeblock.

Step 2) Derive the predictors ψ on the corresponding treeblock of the spatial-temporal, previously coded view, and texture-depth.

Step 3) Perform fast CU depth level range determination. Compute motion complexity parameter *MC* based on equation (5), and classify the current treeblock into static and complex motion region. When the current treeblock is static area, the best depth level range is set to be [0, 1] for texture coding, and "0" for depth coding; when the current treeblock is complex motion area, the best depth level range is set to [2, 3] for texture coding, and [1, 3] for depth coding.

Step 4) Perform early SKIP and Merge mode detection. If the corresponding treeblocks of the current treeblock in predictors ψ are all with the SKIP and Merge mode, skip ME/DE process, and go to Step 6).

Step 5) Perform adaptive termination mode prediction. Compute for *THr* by equations (8) and (9). Compare the current treeblock (texture video or depth map) RD cost with *THr*. When the RD cost of the current treeblock is smaller than *THr*, early terminate the mode decision process.

Step 6) Select the best mode.

IV. EXPERIMENTAL RESULTS

A. EXPERIMENTAL SETUP

To evaluate the performance of proposed overall method, the proposed fast scheme is implemented on recent 3D-HEVC encoder (HTM 16.1 [40]). We have tested eight MVD sequences in two resolutions (1024 × 768 and 1920 × 1088) recommended by JCT-3V Group from common test condition (CTC) [41]. Each test MVD sequence contains three views. In these sequences, the "Shark", "Undo_Dancer", and "GT_Fly" are synthetic video sequences with high precision depth map, the "Kendo", "Balloons", "Newspaper", "Poznan_Hall2", and "Poznan_Street" are natural video sequences with estimated depth map. Test conditions are used as follows: three-view case: center view -left view -right view (in coding order), QP are set as (25, 34), (30, 39), (35, 42) and (40, 45), the maximum coding unit is 64 × 64 pixels, a max partition depth is 4. In this section, we compare the proposed fast method with HTM16.1 and the state-of-the-art fast methods [33], [35], [36], [38], where the compression efficiency is measured by BDBR [42], and "*TS*" means the entire encoding time saving compared with 3D-HEVC, which is defined as follow,

$$TS = \frac{Time_{proposed} - Time_{origional}}{Time_{origional}} \times 100\% \qquad (11)$$

where $Time_{proposed}$ and $Time_{origional}$ denote encoding time of the proposed scheme and original 3D-HEVC for same test video, respectively. In Tables 8-10, "TS_t", "TS_d", and "TS_o" represent run time savings of the texture, depth, and overall coding, respectively. The experimental work is implemented on Windows 7 with two CPU of Intel Xeon E5-2640@2.0 GHz and RAM of 32.0 GB.

B. RESULTS OF THE INDIVIDUAL METHODS

Table 8 gives the individual coding results of the proposed algorithms, i.e., fast CU depth level range determination (FCUDR), early SKIP and Merge mode detection (ESMD), and adaptive early termination mode prediction (AETMP), respectively. Since the overall performance of the proposed individual algorithm has more research value for this paper, the overall results of the proposed individual method are shown in Table 8. For FCUDR strategy, about 46.8% overall runtime has been saved with only 0.42% BDBR increase (or 0.02 dB BDPSNR drop). This encoding time reduction is particularly high for large motion sequences such

TABLE 8. Coding results of each individual method compared with original encoders.

Sequences	FCUDR		ESMD		AETMP	
(%)	BDBR	TS$_o$	BDBR	TS$_o$	BDBR	TS$_o$
Kendo	0.36	-48.7	0.23	-33.2	0.16	-28.1
Balloons	0.29	-46.6	0.15	-31.7	0.12	-27.3
Newspaper	0.57	-42.7	0.28	-29.6	0.19	-24.7
Shark	0.64	-48.9	0.32	-24.8	0.25	-22.8
Undo_Dancer	0.69	-49.2	0.39	-25.1	0.29	-23.2
GT_Fly	0.41	-48.5	0.14	-30.7	0.16	-25.9
Poznan_Street	0.26	-50.2	0.11	-34.6	0.13	-29.3
Poznan_Hall2	0.12	-39.6	0.04	-41.2	0.08	-34.2
Average	0.42	-46.8	0.21	-31.4	0.17	-26.9

TABLE 9. Coding results of the overall algorithm.

Sequences (%)	BDBR	TS$_o$	TS$_t$	TS$_d$
Kendo	0.63	-74.9	-70.3	-78.1
Balloons	0.51	-70.8	-67.5	-73.3
Newspaper	0.96	-63.7	-60.1	-66.9
Shark	1.12	-62.4	-59.7	-65.8
Undo_Dancer	1.37	-62.2	-58.9	-66.1
GT_Fly	0.48	-71.1	-66.9	-76.5
Poznan_Street	0.42	-77.6	-73.7	-79.5
Poznan_Hall2	0.21	-79.2	-75.8	-82.6
Average	0.62	-70.2	-66.6	-73.6

TABLE 10. Results of proposed method and state-of-the art works.

Work	BDBR (%)	TS$_o$ (%)
FDML[32]	0.13	-58.3
FDTFD [33]	0.11	-4.6
FDTC [35]	0.38	-51.6
OLCR [37]	0.33	-22.7
FMDR [38]	1.39	-63.8
FMDG [40]	0.66	-30.6
FMCC [41]	0.59	-60.1
Proposed method	0.62	-70.2

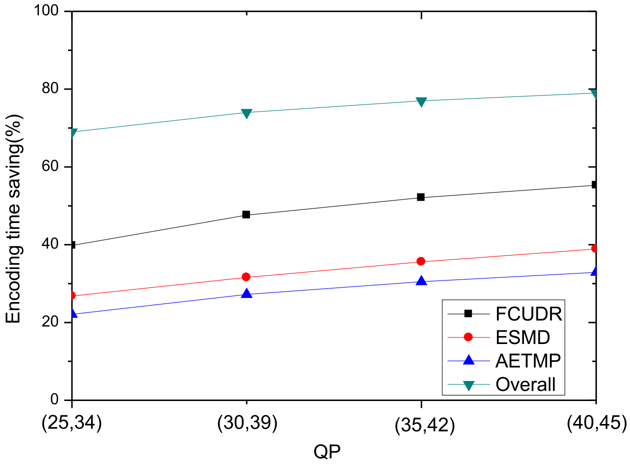
as "Undo_Dancer" (49.2%) and "Shark" (48.9%), but it is still evident for low activity motion sequence such as "Poznan_Hall2" (39.6%). These indicate that FCUDR strategy can skip unnecessary depth levels of 3D-HEVC. For the proposed ESMD approach, about 31.4% overall coding time has been saved, and the bit-rate increase is 0.21%, which is negligible. Therefore, ESMD method can maintain the coding efficiency and greatly reduce the complexity of HTM. In Table 8, the AETMP method shows a consistent gain in overall encoding speed for different motion activities videos with the highest gain of 34.2% in "Poznan_Hall2" and the lowest gain of 22.8% in "Shark". On the contrary, the RD efficiency almost has no loss, only 0.17% bit-rate increase. This result demonstrates that AETMP method can early terminate ME and DE process on unnecessary CU sizes.

C. RESULTS OF OVERALL METHOD

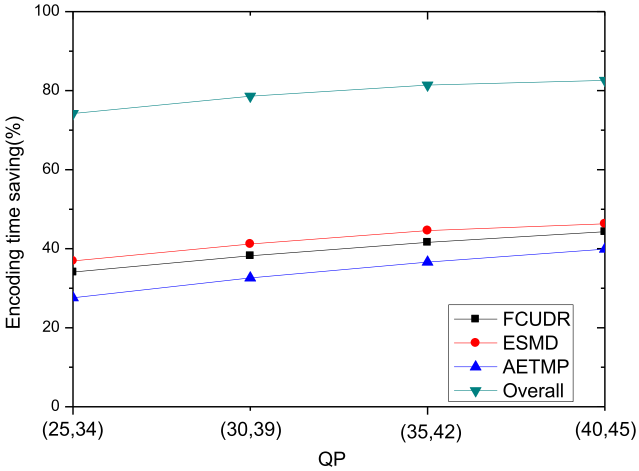
Table 9 gives the coding results of the proposed overall method which includes FCUDR, ESMD and AETMP approach compared with original encoder. The proposed overall method can greatly reduce the encoding time. The proposed overall method can save the coding time by 70.2%. For static or slow motion video like "Poznan_Hall2", the proposed method saves 79.2% overall coding time. The reduction in computation is especially high because exhaustive mode decision process of a large number of treeblocks is not handled for 3D-HEVC. For complex sequence like "Undo_Dancer", the proposed overall method also can save reduce 62.2% overall encoding time. It can be observed that encoding time savings are also constant for 3D-HEVC, which can reduce encoding time by 66.6% and 73.6% in texture and depth, respectively. Meanwhile, the average BD-rate increment is 0.62 % (or BD-PSNR decrease is 0.03dB), which is can be negligible. Therefore, the proposed overall method can efficiently save the encoding time of HTM with negligible loss of RD performance.

Fig. 4 gives the overall time saving of proposed approaches (including FCUDR, ESMD, AETMP and the proposed overall method) compared with HTM encoder for two typical videos "Kendo" and "Poznan_Hall2". The proposed algorithm has consistency for different video sequences. In order to see the performance of individual algorithm and the overall algorithm more clearly, the performance of the sequence is presented separately. As shown in Fig. 4, the FCUDR, ESMD, AETMP and the proposed overall approach attain consistent overall time saving for different resolution test sequences over a large QP values range. Moreover, with QP values increase, the encoding time savings increase in the curves. The reason is that with the QP increase, the probability of depth level 0-1 for texture treeblocks and depth level 0 for depth treeblocks is due to FCUDR approach, the probability of SKIP/Merge mode is due to ESMD approaches, and the probability of early termination mode prediction is due to AETMP approaches, which are all increased in HTM.

In addition to HTM, the proposed overall method is also compared with state-of-the-art works for an objective comparison. These algorithms include FDML [32], FDTFD [33], FDTC [35], OLCR [37], FMDR [38], FMDG [40], and FMCC [41], which are the fast and efficient algorithms



(a) "Kendo"



(b) "Poznan_Hall2"

FIGURE 4. **Overall time saving of the proposed approaches compared with 3D-HEVC.**

for 3D-HEVC. Note that, the FDTC and OLCR method are designed for texture coding only in 3D-HEVC; the FDML and FDTFD method are designed for depth coding only in 3D-HEVC. All the algorithms are used on HTM16.1 and same computer for comparison. Table 10 gives the coding result of the overall time saving and BD-rate, respectively.

We can see from Table 10 that the FDTFD, OLCR, and FMDG algorithm perform good RD performance, but their encoding time savings are poor. Compared with FDTFD, OLCR and FMDG algorithm, the proposed overall method can reduce 39.6%-65.6% encoding time with good RD performance. Among these seven previous methods, FDTFD method has the smallest coding efficiency degradation, and FMDR method achieves the largest computation reduction. Compared with FMDR algorithm, the proposed overall algorithm can increase 6.4% encoding time saving with a better RD efficiency. For state-of-the-art works algorithms, overall encoding time is reduced by 4.6-63.8%, and BDBR increases by 0.11-1.39%. The proposed overall approach achieves 70.2% overall coding time savings and proper RD performance, with only 0.62% increase in BDBR. Compared to FDML, FDTFD, FDTC, OLCR, FMDR, FMDG and FMCC algorithm, the proposed overall method achieves better performance on computation reduction. About 10.1%-65.6% overall coding time is further saved compared with state-of-the-art works. At the same time, the increase in BDBR is negligible. All of the experimental results show that the proposed overall method utilizing the correlations of inter-view, spatio-temporal, and texture-depth is efficient for all categories of test videos and the reduction of the computational complexity outperforms the state-of-the-art fast method for 3D-HEVC.

V. CONCLUSION

In this paper, we presented an efficient algorithm to reduce the encoding time of 3D-HEVC, which involves three works, i.e., fast CU depth level, early SKIP/Merge mode detection, and adaptive early termination mode prediction. It makes use of the correlations of inter-view, spatio-temporal, and texture-depth to skip some treeblocks without a normal coding process. The experimental results demonstrate that the proposed overall method can significantly reduce 70.2% compression time, while it maintains nearly the same RD performances. Furthermore, the proposed overall method outperforms other state-of-the-art approaches for 3D-HEVC with a better complexity reduction.

REFERENCES

[1] Y. Chen and A. Vetro, "Next-generation 3D formats with depth map support," *IEEE MultiMedia*, vol. 21, no. 2, pp. 90–94, Apr./Jun. 2014.

[2] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV," *Proc. SPIE*, vol. 5291, pp. 93–104, May 2004.

[3] G. J. Sullivan, J. M. Boyce, Y. Chen, J.-R. Ohm, C. A. Segall, and A. Vetro, "Standardized extensions of high efficiency video coding (HEVC)," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1001–1016, Dec. 2013.

[4] G. Tech, Y. Chen, K. Muller, J.-R. Ohm, A. Vetro, and Y.-K. Wang, "Overview of the multiview and 3D extensions of high efficiency video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 35–49, Jan. 2016.

[5] K. Müller, H. Schwarz, D. Marpe, C. Bartnik, S. Bosse, H. Brust, T. Hinz, H. Lakshman, P. Merkle, F. Hunn Rhee, G. Tech, M. Winken, and T. Wiegand, "3D high-efficiency video coding for multi-view video and depth data," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3366–3378, Sep. 2013.

[6] Y. Chen, G. Tech, K. Wegner, and S. Yea, "Test model 11 of 3D-HEVC and MV-HEVC," document JCT3V-J1003, Geneva, Switzerland, Feb. 2015.

[7] L. Shen, Z. Liu, T. Yan, Z. Zhang, and P. An, "View-adaptive motion estimation and disparity estimation for low complexity multiview video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 925–930, Jun. 2010.

[8] T. Zhao, S. Kwong, H. Wang, Z. Wang, Z. Pan, and C.-C.-J. Kuo, "Multiview coding mode decision with hybrid optimal stopping model," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1598–1609, Apr. 2013.

[9] J. Lei, J. Sun, Z. Pan, S. Kwong, J. Duan, and C. Hou, "Fast mode decision using inter-view and inter-component correlations for multiview depth video coding," *IEEE Trans. Ind. Informat.*, vol. 11, no. 4, pp. 978–986, Aug. 2015.

[10] C.-H. Yeh, M.-F. Li, M.-J. Chen, M.-C. Chi, X.-X. Huang, and H.-W. Chi, "Fast mode decision algorithm through inter-view rate-distortion prediction for multiview video coding system," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 594–603, Feb. 2014.

[11] W. Zhu, Y. Zheng, P. Chen, and J. Feng, "Fast macroblock encoding algorithm based on rate-distortion activity for multiview video coding," *Signal Process., Image Commun.*, vol. 29, no. 10, pp. 1063–1078, Nov. 2014.

[12] Y. Zhang, S. Kwong, G. Jiang, X. Wang, and M. Yu, ''Statistical early termination model for fast mode decision and reference frame selection in multiview video coding,'' *IEEE Trans. Broadcast.*, vol. 58, no. 1, pp. 10–23, Mar. 2012.

[13] Q. Zhang, P. An, Y. Zhang, L. Shen, and Z. Zhang, ''Low complexity multiview video plus depth coding,'' *IEEE Trans. Consum. Electron.*, vol. 57, no. 4, pp. 1857–1865, Nov. 2011.

[14] J. Yang and A. Wei, ''Fast mode decision algorithm for intra prediction in HEVC,'' in *Proc. IEEE 4th Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, Jun. 2020, pp. 1–4.

[15] T. L. da Silva, L. V. Agostini, and L. A. da Silva Cruz, ''Fast HEVC intra prediction mode decision based on EDGE direction information,'' in *Proc. 20th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2012, pp. 1214–1218.

[16] G. Correa, P. Assuncao, L. Agostini, and L. da Silva Cruz, ''Complexity control of high efficiency video encoders for power-constrained devices,'' *IEEE Trans. Consum. Electron.*, vol. 57, no. 4, pp. 1866–1874, Nov. 2011.

[17] S. Cho and M. Kim, ''Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 9, pp. 1555–1564, Sep. 2013.

[18] L. Shen, Z. Liu, X. Zhang, W. Zhao, and Z. Zhang, ''An effective CU size decision method for HEVC encoders,'' *IEEE Trans. Multimedia*, vol. 15, no. 2, pp. 465–470, Feb. 2013.

[19] L. Shen, Z. Zhang, and P. An, ''Fast CU size decision and mode decision algorithm for HEVC intra coding,'' *IEEE Trans. Consum. Electron.*, vol. 59, no. 1, pp. 207–213, Feb. 2013.

[20] C. Zhou, F. Zhou, and Y. Chen, ''Spatio-temporal correlation-based fast coding unit depth decision for high efficiency video coding,'' *J. Electron. Imag.*, vol. 22, no. 4, Oct. 2013, Art. no. 043001.

[21] T. Zhao, Z. Wang, and S. Kwong, ''Flexible mode selection and complexity allocation in high efficiency video coding,'' *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1135–1144, Dec. 2013.

[22] K. Lee, H.-J. Lee, J. Kim, and Y. Choi, ''A novel algorithm for zero block detection in high efficiency video coding,'' *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1124–1134, Dec. 2013.

[23] J. Xiong, H. Li, Q. Wu, and F. Meng, ''A fast HEVC inter CU selection method based on pyramid motion divergence,'' *IEEE Trans. Multimedia*, vol. 16, no. 2, pp. 559–564, Feb. 2014.

[24] L. Shen, Z. Zhang, and Z. Liu, ''Adaptive inter-mode decision for HEVC jointly utilizing inter-level and spatiotemporal correlations,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 10, pp. 1709–1722, Oct. 2014.

[25] L. Shen, Z. Zhang, and Z. Liu, ''Effective CU size decision for HEVC intracoding,'' *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4232–4241, Oct. 2014.

[26] L. Shen, Z. Zhang, X. Zhang, P. An, and Z. Liu, ''Fast TU size decision algorithm for HEVC encoders using Bayesian theorem detection,'' *Signal Process., Image Commun.*, vol. 32, pp. 121–128, Mar. 2015.

[27] Q. Zhang, J. Zhao, X. Huang, and Y. Gan, ''A fast and efficient coding unit size decision algorithm based on temporal and spatial correlation,'' *Optik*, vol. 126, no. 21, pp. 2793–2798, Nov. 2015.

[28] C.-S. Park, ''Edge-based intramode selection for depth-map coding in 3D-HEVC,'' *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 155–162, Jan. 2015.

[29] H.-B. Zhang, Y.-L. Chan, C.-H. Fu, S.-H. Tsang, and W.-C. Siu, ''Quadtree decision for depth intra coding in 3D-HEVC by good feature,'' in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 1481–1485.

[30] J. Lei, J. Duan, F. Wu, N. Ling, and C. Hou, ''Fast mode decision based on grayscale similarity and inter-view correlation for depth map coding in 3D-HEVC,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 3, pp. 706–718, Mar. 2018.

[31] Q. Zhang, N. Zhang, T. Wei, K. Huang, X. Qian, and Y. Gan, ''Fast depth map mode decision based on depth-texture correlation and edge classification for 3D-HEVC,'' *J. Vis. Commun. Image Represent.*, vol. 45, no. 1, pp. 170–180, May 2017.

[32] M. Saldanha, G. Sanchez, C. Marcon, and L. Agostini, ''Fast 3D-HEVC depth map encoding using machine learning,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 3, pp. 850–861, Mar. 2020, doi: 10.1109/TCSVT.2019.2898122.

[33] H. Hamout and A. Elyousfi, ''Fast depth map intra coding for 3d video compression-based tensor feature extraction and data analysis,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 7, pp. 1933–1945, Jul. 2019, doi: 10.1109/TCSVT.2019.2918770.

[34] N. Zhang, Y.-W. Chen, J.-L. Lin, J. An, K. Zhang, S. Lei, S. Ma, D. Zhao, and W.Gao, *3D-CE3.h Related: Fast Encoder Decision for Texture Coding*, document JCT3V-E0173, ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Vienna, Austria, Jul. 2013.

[35] N. Zhang, D. Zhao, Y.-W. Chen, J.-L. Lin, and W. Gao, ''Fast encoder decision for texture coding in 3D-HEVC,'' *Signal Process., Image Commun.*, vol. 29, no. 9, pp. 951–961, Oct. 2014.

[36] Q. Zhang, X. Wang, X. Huang, R. Su, and Y. Gan, ''Fast mode decision algorithm for 3D-HEVC encoding optimization based on depth information,'' *Digit. Signal Process.*, vol. 44, pp. 37–46, Sep. 2015.

[37] H. R. Tohidypour, M. T. Pourazad, and P. Nasiopoulos, ''Online-learning-based complexity reduction scheme for 3D-HEVC,'' *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 10, pp. 1870–1883, Oct. 2016.

[38] L. Shen, P. An, Z. Zhang, Q. Hu, and Z. Chen, ''A 3D-HEVC fast mode decision algorithm for real-time applications,'' *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 11, no. 3, p. 34, Jan. 2015.

[39] L. Shen, P. An, and Z. Liu, ''Context-adaptive based CU processing for 3D-HEVC,'' *PLoS ONE*, vol. 12, no. 2, Feb. 2017, Art. no. e0171018.

[40] Q. Zhang, R. Jing, B. Wang, P. Cui, C. Zhou, T. Yan, S. Kanyukt, and W. Si, ''Fast mode decision based on gradient information in 3D-HEVC,'' *IEEE Access*, vol. 7, pp. 135448–135456, Sep. 2019, doi: 10.1109/ACCESS.2019.2940706.

[41] Q. Zhang, K. Huang, X. Wang, B. Jiang, and Y. Gan, ''Efficient multiview video plus depth coding for 3D-HEVC based on complexity classification of the treeblock,'' *J. Real-Time Image Process.*, vol. 16, no. 6, pp. 1909–1926, Dec. 2019, doi: 10.1007/s11554-017-0692-5.

[42] Q. Zhang, H. Chang, X. Huang, L. Huang, R. Su, and Y. Gan, ''Adaptive early termination mode decision for 3D-HEVC using inter-view and spatio-temporal correlations,'' *AEU-Int. J. Electron. Commun.*, vol. 70, no. 5, pp. 727–737, May 2016.

[43] (Dec. 2017). *3D-HEVC Software HTM-16.1*. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn3DVCSoftware/tags/HTM-16.1

[44] K. Mueller and A. Vetro, *Common Test Conditions of 3DV Core Experiments*, document JCT3V-G1100, Joint Collaborative Team on 3D Video Coding Extensions (JCT-3V) 7th Meeting, San Jose, CA, USA, Jan. 2014.

[45] G. Bjontegaard, *Calculation of Average PSNR Differences Between RD Curves*, document ITU-T SG16 Q.6, VCEG Meeting, VCEG-M33, Austin, TX, USA, Apr. 2001.

QIUWEN ZHANG (Member, IEEE) received the Ph.D. degree in communication and information systems from Shanghai University, Shanghai, China, in 2012. Since 2012, he has been a Faculty Member with the College of Computer and Communication Engineering, Zhengzhou University of Light Industry, where he is currently an Associate Professor. He has published over 30 technical articles in the field of pattern recognition and image processing. His major research interests include 3D signal processing, machine learning, pattern recognition, video codec optimization, and multimedia communication.

YIHAN WANG received the B.S. degree in information engineering from the Zhengzhou University of Light Industry, Zhengzhou, China, in 2018, where she is currently pursuing the master's degree in information and communication systems with the School of Computer and Communication Engineering. Her current research interests include image processing, 3D video coding, and extensions of the high efficiency video coding.

TAO WEI received the B.S. degree in computer science and technology from the PLA Information Engineering University, China, in 2003, the M.S. degree in computer science and technology from the Nanjing University of Science and Technology, China, in 2010, and the Ph.D. degree in software engineering from China National Digital Switching System Engineering & Technological Research and Development Center, China, in 2019. He is currently an Associate Professor with the Computer College, Henan Institute of Engineering, China. His research interests include image processing, video codec optimization, and multimedia communication.

LIXUN HUANG received the B.S. degree in electronic information science and technology and the M.S. degree in computer application technology from the Henan University of Technology, China, in 2006 and 2009, respectively, and the Ph.D. degree in communication and information system from Shanghai University, China, in 2013. He has been a Lecturer with the Zhengzhou University of Light Industry, since July 2013. His research interests include iterative learning control and pattern recognition.

RIJIAN SU received the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2010. He is currently a Professor with the Zhengzhou University of Light Industry, China. His current research interests include image processing, Bayesian estimation, and intelligent computation.

• • •