Fast Depth Map Coding Based on Bayesian Decision Theorem for 3D-HEVC

DONGYAO ZOU, PU DAI[ID]**, AND QIUWEN ZHANG**[ID]
College of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China

Corresponding author: Qiuwen Zhang (2012032@zzuli.edu.cn)

ABSTRACT The depth map compression of 3D High-Efficiency Video Coding (3D-HEVC) inherits prediction structure adopted by HEVC and develops supplementary coding modes to better express depth image. These new coding modes combine with existing technology implement high coding efficiency, but lead to an extremely huge increase in coding time. This article proposes a rapid coding method for depth map in 3D-HEVC. The proposed scheme utilizes the Bayesian decision rule and the correlations between corresponding texture video and spatially adjacent treeblocks to analyze the treeblock features of depth map. According to the analysis, we propose two approaches, including early SKIP/Merge mode selection, and adaptive CU pruning termination. Simulation consequences illustrate that this paper can save 51.2% complexity while maintaining the rate distortion (RD) performance.

INDEX TERMS 3D-HEVC, depth video, complexity, Bayesian.

I. INTRODUCTION

During the last few years, three-dimensional (3D) video applications, such as 3D movies, 3D games, and FTV, have attracted a greatly attention [1]. MPEG is currently developing mulit-view video plus depth (MVD), which is depth-enhanced 3D representation [2], [3]. The depth map denotes the geometric distance between the objects and camera that can generate intermediate views by utilizing the depth image-based rendering (DIBR) [4]. Therefore, JCT-3V exploit 3D-HEVC for enhancing the denseness performance of depth map [5], [6].

Since depth image has large homogeneous areas demarcated by sharp object edges, there is a difference between the characteristic of depth map and texture video [7]. Different from texture coding, it is significant to conserve the depth sharp edge compared with the visual quality. According to the feature, some prediction techniques are designed in 3D-HEVC, such as Depth Modeling Modes (DMMs) [8], Depth Intra Skip [9], and segment-wise depth coding (SDC) [10] for a better compression of edges in depth image. The increase in coding performance is accompanied by a substantial increase in coding time. Consequently,

a rapid depth coding approach is devised in this article, which decrease the coding runtime with maintaining coding quality.

A number of literatures about fast depth coding methods [11]–[17] have been reported to decrease the complexity of multi-view video coding (MVC). All these algorithms show good performance in terms of reducing coding complexity, while loss of video quality for MVC is very small. However, these approaches are not well appropriate for 3D-HEVC, where new additional quadtree structured and depth intra modes introduced for 3D-HEVC are not taken.

Recently, studies in 3D-HEVC have been conducted to decrease the depth complicacy. In [18], a fast method for depth map of 3D-HEVC is introduced for allocating the complexity, in which used the edge classification for the intra prediction selectively omits DMMs. In [19], a fast coding method is devised to speed up the depth map prediction processes, where it employs the correlation of neighboring CUs depth and the correlation of texture-depth to adaptively adjust the mode decision processing. Based on an adaptive threshold pattern, an early SKIP is developed in [20] to decrease complicacy of the inter frames. In [21], a fast approach is utilized based on detecting edge regions to increase depth coding runtime saving, when the treeblock does not contain object edges in depth map, the DMMs are not tested. In [22], a fast approach and early mode decision method are presented

to expedite the depth compression. Reference [23] devices a fast depth coding approach to decrease the number of modes. A quadtree limitation method based on data mining is introduced in [24] for reducing the complexity of depth map. A low complexity intra decision approach is utilized based on the correlations of spatiotemporal, inter-component and inter-view in [25] for accelerating coding. Reference [26] designs a fast edge detection-based approach for the depth map, which increases the time saving of depth map while maintaining the encoding efficiency. By discovering the flat region and texture orientation of the depth map, [27] devise a fast edge detection-based decision method that can decrease the encoding runtime. Reference [28] introduces fast static decision trees-based coding method for the depth map of 3D-HEVC. Fast approaches for the mode decision of 3D-HEVC are devised in our previous works [29] and [30] to reduce the coding complexity. The above methods for 3D-HEVC are well devised depth compression with achieving coding time saving. Nevertheless, most of the above schemes significantly reduce the RD performance.

Bayesian theorem that is a promising method has been proved to be able to effectively reduce the coding complexity, so it is widely used in video compression applications. In some recent works of HEVC, Bayesian has been researched [31]–[34]. To the best of our knowledge, there are few studies on its depth map compression using Bayesian theorem in 3D-HEVC. This paper designs the depth map coding of 3D-HEVC based on the fast Bayesian theorem, which composes early SKIP/Merge mode decision and adaptive CU pruning termination.

The remaining of this paper is divided as follows: The mode statistical analysis of depth map coding is introduced in Section 2. Section 3 describes the proposed rapid method to proceed in detail. Sections 4 discusses experimental results and Sections 5 considers conclusion.

II. OBSERVATIONS AND ANALYSIS

At all the possible depth levels in depth coding of 3D-HEVC, the optimum mode is the mode with the smallest RD cost in the mode decision method, which is performed as,

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

where SSE_{luma} and SSE_{chroma} are the distortion of the present treeblock on luminance and its rehabilitated treeblock on chromaticity, respectively. ω_{chroma} denotes the weight parameter in chroma; λ_{mode} and R_{mode} are Lagrange multiplier and the total bitrate cost, respectively. The "try all then select the best" scheme provides the excellent RD performance, however also requires heavy complicacy which restricts encoder from the real-time application.

To better understand the features of depth map, some statistical researches are carried out, including the prediction 3D-HEVC coder pattern and the allocation of calculation complicacy. The experiment is tested on 3D-HEVC

(HTM 16.1 [35]) by the JCT-3V. Among these videos, the 1920 × 1088 and 1024 × 768 resolution sequences include "Undo_Dancer", "GT_Fly", "Poznan_Street", "Poznan_Hall2", "Shark", "Kendo", "Balloons" and "Newspaper". The simulation settings include: 3-view, quantization parameters (QPs): 20, 25, 35 and 45; VSO is on; DMMs are on; 150 frames are conducted simulation.

A. INTER PREDICTION MODE DISTRIBUTION

Table 1 shows that an inter mode distribution among different QPs are studied by utilizing the HTM 16.1. It can be seen that most treeblocks select SKIP/Merge as the optimal mode in inter coding, the percentage of SKIP/Merge treeblocks increases with the increase of QP. About 87.1% select SKIP, while no more than 5.1% select Merge, 2N×2N, Intra, and

TABLE 1. Inter mode distribution.

Sequences	QP	SKIP/ Merge	Inter 2N×2N	Intra modes	Other inter modes
Kendo	20	62.8%	12.9%	15.8%	8.5%
	25	85.8%	5.2%	3.2%	5.8%
	35	93.1%	2.9%	1.6%	2.4%
	45	96.5%	1.6%	1.1%	0.8%
Balloons	20	69.9%	9.1%	9.3%	11.7%
	25	88.7%	4.6%	3.9%	2.8%
	35	94.5%	2.2%	1.9%	1.4%
	45	98.1%	0.8%	0.6%	0.5%
Newspaper	20	63.8%	9.2%	8.1%	18.9%
	25	85.2%	5.4%	4.2%	5.2%
	35	93.5%	3.1%	1.9%	1.5%
	45	96.4%	1.2%	1.1%	1.3%
Shark	20	64.9%	8.1%	7.3%	19.7%
	25	85.7%	3.3%	3.2%	7.8%
	35	93.7%	1.2%	1.4%	3.7%
	45	96.2%	0.6%	0.7%	2.5%
Undo_Dancer	20	67.1%	7.9%	7.1%	17.9%
	25	86.4%	4.0%	2.6%	7.0%
	35	94.7%	1.5%	1.1%	2.7%
	45	96.2%	0.9%	0.6%	2.3%
GT_Fly	20	73.9%	6.1%	11.7%	8.3%
	25	89.6%	3.5%	4.9%	2.0%
	35	94.3%	2.1%	2.6%	1.0%
	45	98.2%	0.7%	0.6%	0.5%
Poznan_Street	20	72.4%	6.7%	9.6%	11.3%
	25	88.1%	3.6%	4.7%	3.6%
	35	94.0%	2.1%	2.3%	1.6%
	45	98.2%	0.6%	0.5%	0.7%
Poznan_Hall2	20	76.7%	6.2%	9.3%	7.8%
	25	92.8%	2.5%	2.7%	2.0%
	35	96.9%	1.2%	0.9%	1.0%
	45	98.9%	0.3%	0.4%	0.4%
Average		87.1%	3.8%	4.0%	5.1%

other inter modes in depth map treeblock. Since depth videos with more fixed area mostly chooses the SKIP/Merge mode. Particular for small global motion sequence selecting SKIP mode is extremely high, that is 98.9% in QP 45. Therefore, if the proposed scheme pre-determined SKIP/Merge as the best inter mode of depth map, lots of computation can be reduced.

B. CU DEPTH LEVEL DISTRIBUTION

Table 2 illustrates the depth level distribution. The CU depth levels of depth map includes "Level 0", "Level 1", "Level 2" and "Level3". It is observed from Table 2 that 66.5%, 17.9%, 9.2% and 6.5% of depth map treeblocks select depth "Level 0", "Level 1", "Level 2", and "Level 3", respectively. The percentage of selecting "Level 0" is considerably high on the average value, such as sequence "Poznanhall2" with a large homogeneous texture is 78.6%. Simultaneously, it is found that about 84.4% of treeblocks select the "Level 0" and "Level 1", and selecting "Level 2" and "Level 3" as the best depth level is only about 15.7% of depth treeblocks.

Fig.1 shows the complexity in different maximum CU depth levels. From the figure we can conclude that coding complexity increases considerably with the increase of depth. In addition, Table 2 demonstrates that the percentage of a depth map treeblock coded with "Level 2" and "Level 3" are very small. During CU pruning procedure, if we utilize an adaptive termination scheme, the coding time saving is increased in a larger CU depth, but the RD performance will slightly degrade. Thus, if an early termination method is determined in CU pruning procedure, the increase of compression time will be ignored in larger CU depth.

TABLE 2. CU depth level distribution of depth map coding.

Sequences	Level0 (%)	Level1 (%)	Level2 (%)	Level3 (%)
Kendo	67.3	16.8	9.1	6.8
Balloons	69.2	15.6	8.7	6.5
Newspaper	64.2	18.8	9.8	7.2
Shark	58.4	23.1	11.2	7.3
Undo_Dancer	59.1	21.4	10.7	8.8
GT_Fly	68.2	16.5	8.5	6.8
Poznan_Street	66.8	17.5	9.5	6.2
Poznan_Hall2	78.6	13.2	6.1	2.1

III. PROPOSED ALGORITHM

A. EARIL SKIP/MERGE MODE SELECTION BAESD ON BAYESIAN DECISION RULE

In depth compression of 3D-HEVC, the inter treeblocks are selected including SKIP/Merge that has no reference to picture index and coded motion vector delta in encoding process, normal inter, and intra modes. Thus, when SKIP/Merge is pre-decided, the mutable inter mode calculation of the current CU in depth coding is notably decreased. As analyzed in Section 2, if depth map treeblock is early discerned by
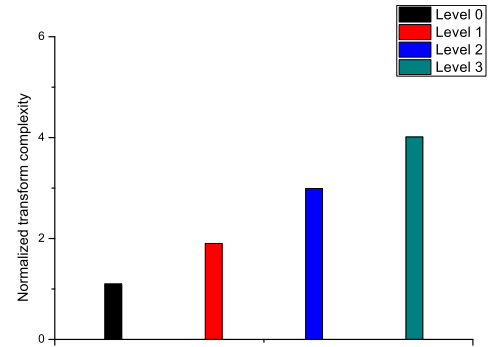

FIGURE 1. Complexity in different maximum CU depth levels.

looking for an appropriate condition, then SKIP/Merge mode will be chosen probably. In 3D-HEVC encoders, it can simplify the inter mode decision and skip large computational complexity.

Bayesian classifier is a formula that computes conditional probability and calculate posterior probability given several observations, and it also can detect if the inter mode for the current treeblock is optimal. According to this concept, an early SKIP/Merge decision based on Bayesian theorem is developed before the full RD calculating for depth coding. If SKIP/Merge is selected based on Bayesian decision rule, then the current treeblock is coded as SKIP/Merge mode, and other inter-intra modes are no longer calculated. Meanwhile, if the Bayesian determines the current treeblock which is not SKIP/Merge, and then other inter-intra modes are tested by the 3D-HEVC encoder.

The SKIP/Merge probability depends on the depth video itself. The SKIP/Merge distribution of depth map is given in Section 2. It represents the SKIP/Merge probability for different QPs in depth coding. Now, Table 1 shows that the percentage of SKIP/Merge varies markedly by the QP value increase, and selecting SKIP/Merge mode is more frequent than the other inter modes. According to this characteristic, early SKIP/Merge selection based on Bayesian decision rule is introduced to avoid testing unnecessary inter modes of the 3D-HEVC depth map.

FIGURE 2. Predictors of the treeblock in depth view.

Let κ_1, κ_2, κ_3, κ_4 denote SKIP/Merge, Inter 2N×2N, Intra modes, and other inter modes, respectively. For determining the SKIP/Merge mode, Bayesian theorem is utilized to obtain

the posterior probability $P\left(\kappa_i/x\right)$,

$$P\left(\kappa_i/x\right) = \frac{P\left(\kappa_i\right) \cdot p\left(x/\kappa_i\right)}{p\left(x\right)} \qquad (2)$$

where $p\left(x\right)$ denotes the mixture density function, $P\left(\kappa_i\right)$ and $p\left(x/\kappa_i\right)$ denote a priori probability of κ_i and the conditional probability density function, respectively.

According to Bayesian decision theorem, that coding mode is defined by,

$$P\left(\kappa_i\right) \cdot p\left(x/\kappa_i\right) > P\left(\kappa_j\right) \cdot p\left(x/\kappa_j\right) \qquad (3)$$

where $P\left(\kappa_i\right)$ and $p\left(x/\kappa_i\right)$ are unknown in depth encoding, we can estimate them from the corresponding texture treeblocks (as shown in Fig.2). Simulation results verified that the a priori probabilities are similar in corresponding texture treeblocks [30]. Thus, $P\left(\kappa_i\right)$ and $p\left(x/\kappa_i\right)$ from the corresponding texture treeblocks are estimated. According to this content, a tolerance threshold ε is introduced to choose the optimal inter mode κ_i' for depth coding,

$$\frac{P\left(\kappa_i'\right) \cdot p\left(x/\kappa_i'\right) - P\left(\kappa_i\right) \cdot p\left(x/\kappa_i\right)}{\sum\limits_{j=1}^{4} P\left(\kappa_j\right) \cdot p\left(x/\kappa_j\right)} \leq \varepsilon \qquad (4)$$

The inter mode prediction will stop if this condition is met, and determine the SKIP/Merge mode for the current depth treeblock early.

B. ADAPTIVE CU PRUNING TERMINATION BASED ON BAYESIAN DECISION RULE

In Section 3.2, the Bayesian decision rule allows for an early SKIP/Merge mode selection. If the Bayesian decision rule is skipped, then the current depth treeblock is coded as SKIP/Merge pattern and other depth levels are not computed. On the other hand, if the Bayesian rule determines SKIP/Merge mode as non-effective, the CU should be split. Therefore, an adaptive CU pruning termination method for depth treeblock based on Bayesian classifier is exploited to further decrease the computational complexity.

Each CU of depth map has two conditions, not-splitting and splitting. Let η_1 and η_2 represent the pattern of not-splitting and splitting, respectively. For a CU of depth map, we use Bayesian theorem for the adaptive CU pruning termination,

$$P\left(\eta_i/y\right) = \frac{P\left(\eta_i\right) \cdot p\left(y/\eta_i\right)}{p\left(y\right)} \qquad (5)$$

where $P\left(\eta_i/y\right)$ denotes a posteriori probability of choosing pattern η_i based on an observation y. $P\left(\eta_i\right)$ denotes a priori probability, $P\left(\eta_i/y\right)$ and $p\left(y\right)$ denote the probability density and the mixture density function, respectively.

In fact, the prediction modes of the depth map treeblock and the corresponding texture and spatially adjacent treeblocks are similar. Based on this analysis, a mode complexity parameter (DMC) is defined as,

$$DMC = \sum\limits_{\ell} \sigma_i \cdot \varsigma_i \qquad (6)$$

where σ denotes the weight factor, $\sum\limits_i \sigma_i = 1$; ς denotes the directions weight factor. ς is defined based on the impact of the relevant treeblock of the current depth treeblock; ℓ is a set of predictors including the spatial and co-located texture video treeblocks. Our method uses the DMC in a depth treeblock as the observed feature on Bayesian decision rule. It can be observed that most of not-splitting CUs are with small value of DMC. At the same time, the percentage of not-splitting CUs gradually obtain smaller value of DMC. Based on the extensive experiments, when the value of DMC is larger than 1.2, the $p\left(y/\eta_2\right)$ is larger than $p\left(y/\eta_1\right)$.

In 3D-HEVC, making the wrong decision will cause coding efficiency loss. The proposed algorithm uses CL_1 which represents the loss of compression efficiency for a splitting CU wrong choice as the not-splitting CU. Meanwhile, CL_2 is represented as the loss of compression efficiency for a not-splitting CU wrong choice as the splitting CU. Based on the above explanation, Bayesian cost$BC\left(\eta_i/y\right)$ is calculated as follow,

$$\begin{cases} BC\left(\eta_1/y\right) = p\left(m_1\right) \cdot p\left(x/\eta_1\right) \cdot E\left[CL_1\right]/p\left(y\right) \\ BC\left(\eta_2/y\right) = p\left(m_2\right) \cdot p\left(x/\eta_2\right) \cdot E\left[CL_2\right]/p\left(y\right) \end{cases} \qquad (7)$$

where $E\left[\cdot\right]$ represents the expected value.

When $BC\left(\eta_1/y\right)$ is smaller than $BC\left(\eta_2/y\right)$, it means that mode η_1 should be choose and we can run early CU pruning termination procedure. To explain the Bayesian error, $p\left(\eta_i\right)$ and $p\left(y/\eta_i\right)$ are predicted from previously encoded depth map, we present the tolerance parameter χ,

$$\frac{p\left(m_2\right) \cdot p\left(x/\eta_2\right) \cdot E\left[CL_2\right]/p\left(y\right)}{p\left(m_1\right) \cdot p\left(x/\eta_1\right) \cdot E\left[CL_1\right]/p\left(y\right)} \leq \chi \qquad (8)$$

where the tolerance parameter range is $\chi \leq 1$. When a treeblock of depth map be in line with Eq. (8), the processing of the CU pruning can be terminated. In this section, $p\left(\eta_i\right)$ and $p\left(y/\eta_i\right)$ are not clear in the current depth treeblock, we can access these parameters from previously coded frames of depth map. Therefore, the proposed method does not require a different training sequence. Actually, the proposed method uses a few depth map frames (about 5 frames) to compute the primary values for the prior and likelihood. Therefore, Bayesian decision rule does not provide any select in the first training frame of the depth video. Moreover, the full RDO of the HTM encoder supplies all the determinations to compute the appropriate parameters. And the proposed scheme is used for the successive frames of depth map.

C. OVERALL ALGORITHM

In the light of the analysis in the previous section, the overall method consists of early SKIP/Merge mode selection and adaptive CU pruning termination. The specific steps of the overall proposed method are as below:

Step 1: Begin execution mode decision with a depth treeblock.

Step 2: Perform early SKIP/Merge selection. Compute the a priori probabilities

$$P(v_i) \, and \, p(x/v_i)$$

from the corresponding texture treeblocks.

Step 3: Determine the RD cost the current depth treeblock based on Eq. (4), if

$$\frac{P(\kappa_i') \cdot p(x/\kappa_i') - P(\kappa_i) \cdot p(x/\kappa_i)}{\sum\limits_{j=1}^{4} P(\kappa_j) \cdot p(x/\kappa_j)} \le \varepsilon,$$

merely SKIP/Merge utilized for optimal mode, and get to Step 6.

Step 4: Execute adaptive CU pruning termination. Calculate the $p(\eta_i)$ and $p(y/\eta_i)$ from previously coded frames of depth map. Compute the DMC of current treeblock in Eq. (6), look up the $p(y/\eta_i)$ based on DMC.

Step 5: If

$$\frac{p(m_2) \cdot p(x/\eta_2) \cdot E[CL_2]/p(y)}{p(m_1) \cdot p(x/\eta_1) \cdot E[CL_1]/p(y)} \le \chi$$

terminate the CU pruning processing, else the current treeblock of depth map is split.

Step 6: Determine the optimum mode.

IV. RESULTS AND ANALYSIS

In order to accurately assess the performance of the suggested approach, eight MVD videos in CTC [35] are utilized to conduct the experiment on HTM 16.1 [36]. Test sequences mainly include the composite video with high exactness depth map and the natural video with estimated depth map in Table 3. The simulation settings are as follows: three-view case, QPs: (25, 34), (30, 39), (35, 42) and (40, 45). The "VSRS-1D-Fast" software is utilized for tests [37]. The Bjontegaard Delta Bitrate (BDBR) [38] is utilized to evaluate the coding performance of the suggested algorithm. The coding runtime saving "*TS*" is employed to measure the complexity, which is defined as,

$$TS = \frac{Time_{proposed} - Time_{Anchor}}{Time_{Anchor}} \times 100\% \qquad (9)$$

where $Time_{proposed}$ and $Time_{Anchor}$ represent the coding runtime saving of the suggested algorithm and anchor approach, respectively.

A. RESULTS OF INDIVIDUAL STAGE

Tables 3-4 illustrate the results of the suggested algorithm, which includes early SKIP/Merge mode selection (ESMMS), and adaptive CU pruning termination (ACUPT). Table 3 illustrates the performance of the ESMMS approach with 3-different ε values compared with HTM16.1. It is observed from Table 3 that ESMMS approach can extremely decrease depth encoding time with resemble coding performance. It saves the depth encoding time by 37.6%,

TABLE 3. Results of the ESMMS compared with HTM16.1.

Sequences	$\varepsilon=0.1$		$\varepsilon=0.5$		$\varepsilon=0.9$	
(%)	BDBR	TS	BDBR	TS	BDBR	TS
Kendo	3.31	-54.3	1.01	-47.6	0.37	-38.9
Balloons	3.09	-52.8	0.96	-46.5	0.26	-37.6
Newspaper	4.13	-49.2	1.11	-44.9	0.43	-35.1
Shark	5.23	-45.1	1.32	-39.8	0.57	-33.3
Undo_Dancer	4.98	-44.6	1.29	-38.7	0.52	-32.8
GT_Fly	3.12	-55.4	0.84	-48.1	0.32	-39.3
Poznan_Street	2.75	-56.2	0.73	-49.2	0.25	-40.8
Poznan_Hall2	2.27	-59.2	0.61	-52.4	0.19	-43.1
Average	3.61	-52.1	0.98	-45.9	0.36	-37.6

TABLE 4. Results of the ACUPT compared with HTM16.1.

Sequences	$\chi = 0.2$		$\chi = 0.5$		$\chi = 0.8$	
(%)	BDBR	TS	BDBR	TS	BDBR	TS
Kendo	0.07	-11.2	0.16	-13.4	0.25	-14.6
Balloons	0.06	-10.7	0.14	-12.2	0.21	-13.9
Newspaper	0.12	-10.1	0.21	-11.4	0.32	-12.7
Shark	0.17	-9.2	0.28	-10.3	0.41	-11.2
Undo_Dancer	0.15	-8.9	0.26	-10.2	0.39	-10.9
GT_Fly	0.05	-11.5	0.13	-13.6	0.23	-14.9
Poznan_Street	0.05	-12.2	0.11	-14.1	0.18	-15.1
Poznan_Hall2	0.03	-13.7	0.09	-15.2	0.13	-16.5
Average	0.09	-10.9	0.17	-12.5	0.27	-13.7

45.9% and 52.1% when $\varepsilon = 0.1$, 0.5, and 0.9, respectively. The large homogeneous sequence, such as "Poznan_Hall2", can extremely reduce coding runtime, but it is still distinct for large motion sequence, such as "Undo_Dancer". Meanwhile, the loss of RD performance is acceptable, less 1.16% BDBR increase on synthesized view when $\varepsilon = 0.5$, and 0.9, respectively. The simulation results demonstrate that ESMMS method can efficiently skip unnecessary inter mode. Table 4 shows that about 10.9%, 12.5%, and 13.7% encoding time have been saved on average when $\chi = 0.2$, 0.5, and 0.8, respectively. However, the average BDBR of the synthesized view increase is less 0.27% with the ignored loss of coding efficiency. Thus, the ACUPT approach can maintain the coding performance and reduce the CU processing time.

B. THE PERFORMANCE OF THE PROPOSED OVERALL SCHEME

Table 5 indicates the results of the overall scheme that includes ESMMS and ACUPT method. According to the simulation results in Section 4.1, we set the $\varepsilon = 0.5$, and $\chi = 0.8$. This is because that the RD performance and encoder complexity are rebalanced by setting two parameters. It can be found from Table 5 that the encoding runtime saving of the overall scheme is about 51.2%, where sequence "Poznan_Hall2" has the maximum 57.1% and

sequence "Undo_Dancer" has minimum 45.9%. Compared with other test sequences, the time reduction of sequence "Poznan_Hall2" is larger, because "Poznan_Hall2" contains more homogeneous regions. Meanwhile, the average BDBR increment of the synthesized view is 1.07%. Thus, the overall scheme can greatly decrease the encoding runtime with the similar RD performance as the original HTM16.1.

Fig.3 illustrates more detailed results of the entire approach for test sequence "Kendo" and "Undo_Dancer". It is found from Fig.3 that the overall scheme has consistent encoding runtime saving while maintaining the similar RD performance. In addition, the coding time savings will increase when the compression bitrate increase and QP reduce. Specifically, the percentage of SKIP/Merge is due to ESMMS. With the increment of QP, the percentage of not-splitting for ACUPT are both increased.

C. COMPLEXITY ANALYSIS

As depicted in Section 3.2, the coding parameters of the density distribution are first offline trained, and then the parameters are updated by using several frames (about 5 frames). Therefore, only the compression time of the training frame will influence the complexity. Based on this, the ratio of training time *OT* is described as follows,

$$OT = \frac{T_t}{T_t + 29 \cdot T_o} \tag{10}$$

where T_t and T_o denote the coding time of training frame and other testing frames, respectively. Fig.4 shows the proportion time of *OT* with the proposed scheme. It can observe that the training frame spends 4.4% of entire compression time on average, which can be acceptable ratio. Furthermore, when the computational complexity is decreased, the encoding runtime consumed by the training frame is included in the coding results of the entire algorithm. Simulation consequences from Fig.4 show that the suggested method can be used in practical applications.

D. RESULTS OF THE PROPOSED SCHEME COMPARED WITH THE STATE-OF-THE-ART METHODS

Except for HTM16.1 encoder, the suggested entire scheme is also compared to the well-known rapid and efficient 3D-HEVC methods that includes EBIMS [18], CRDSD [20], FMDGI [23], EISSI [25], and FMDTE [30] in Figs. 5 and 6. Table 5 shows that the coding time saving of the overall method is 51.2% and the BDBR increases only 1.07% for synthesized view. We can see from Figs. 5 and 6 that five recent methods have fine coding performance, yet the encoding runtime savings of the last methods are less than the proposed scheme. Specifically, the BDBR of the EISSI scheme is the smallest, while the BDBR of the FMDGI approach is the largest in these methods. For all the 3D sequences, the proposed overall algorithm exhibits the largest computation decrease. Compared with five recent methods, the proposed
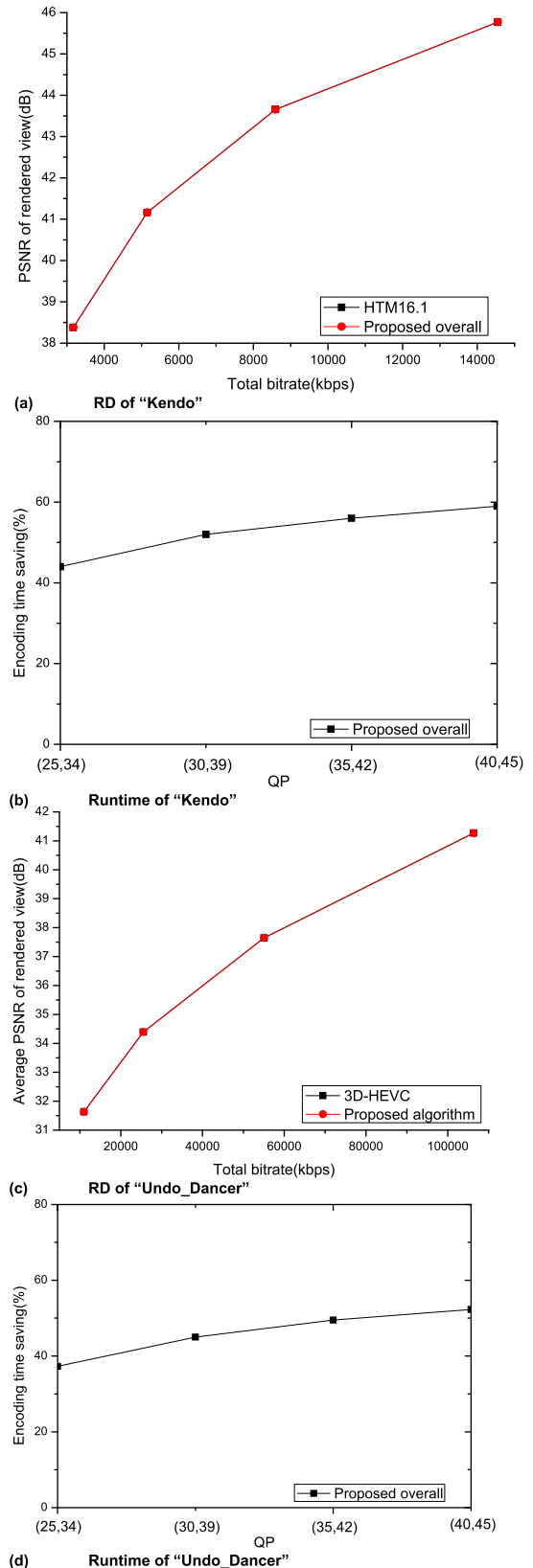

FIGURE 3. RD property and depth time saving compared with HTM16.1.

overall method has superior computation decrease. About 8.0%-29.7% encoding time of the depth coding is further

saved. Moreover, the BDBR loss can be negligible, less than 0.7% BDBR increase compared with EBIMS CRDSD, EISSI, and FMDTE, and with a better coding efficiency than FMDGI method. Compared with five recent methods, the proposed overall scheme has best encoding time saving with almost equal or better RD performance. Simulation results show that the overall scheme can surpass the latest scheme and maintain a better coding performance.

TABLE 5. The performance of the overall method.

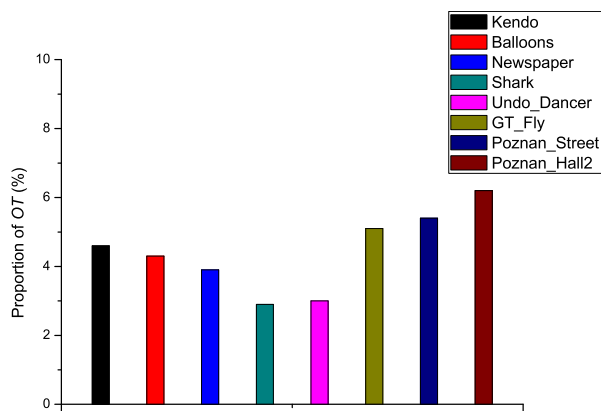| Sequences	$\varepsilon = 0.5, \chi = 0.8$	
(%)	BDBR	TS
Kendo	1.09	-52.2
Balloons	1.02	-51.7
Newspaper	1.21	-49.8
Shark	1.47	-46.5
Undo_Dancer	1.36	-45.9
GT_Fly	0.93	-52.8
Poznan_Street	0.85	-53.4
Poznan_Hall2	0.65	-57.1
Average	1.07	-51.2

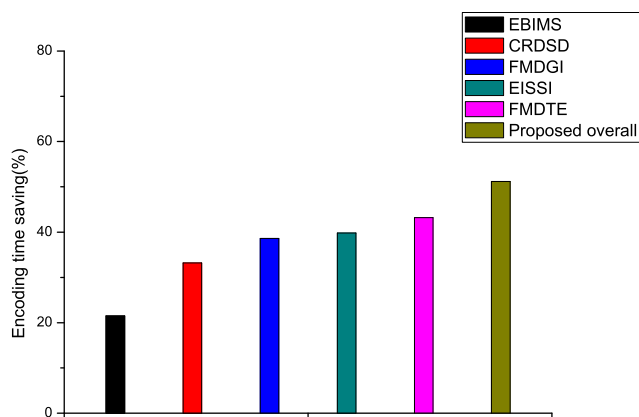**FIGURE 4.** Complexity analysis of the online-training time.

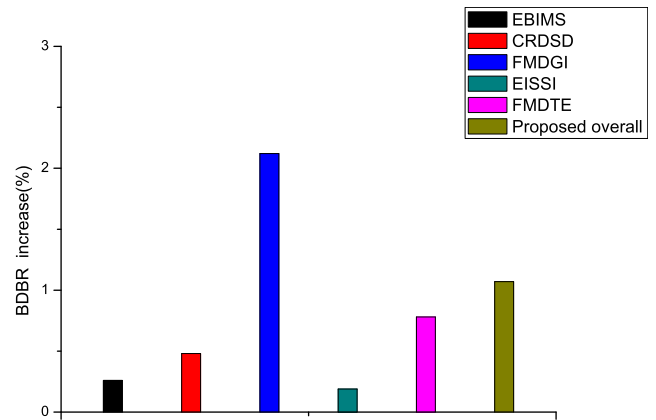**FIGURE 5.** Time saving of proposed approach and five recent methods.

FIGURE 6. BDBR increase of proposed scheme and five recent methods.

V. CONCLUSION

To decrease encoding time, a fast scheme is introduced for the depth map compression of 3D-HEVC in this article, which consists of the early SKIP/Merge mode selection scheme and adaptive CU pruning termination scheme. The proposed scheme mainly utilizes the Bayesian decision rule and the correlations between corresponding texture video and spatially adjacent treeblocks to analyze the treeblock properties of depth map and avoid the calculation of the unnecessary modes. Simulation consequences indicate that the proposed algorithm can effectively decrease coding complexity with only an ignorable loss of RD performance. Furthermore, the computational complexity and coding efficiency are rebalanced by regulating the parameters.

REFERENCES

[1] O. Stankiewicz, M. Domanski, A. Dziembowski, A. Grzelka, D. Mieloch, and J. Samelak, "A free-viewpoint television system for horizontal virtual navigation," *IEEE Trans. Multimedia*, vol. 20, no. 8, pp. 2182–2195, Aug. 2018.

[2] K. Müller, P. Merkle, and T. Wiegand, "3-D video representation using depth maps," *Proc. IEEE*, vol. 99, no. 4, pp. 643–656, Apr. 2011.

[3] 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.

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

[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] G. Tech, Y. Chen, K. Müller, 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.

[7] Q. Zhang, N. Li, L. Huang, and Y. Gan, "Effective early termination algorithm for depth map intra coding in 3D-HEVC," *Electron. Lett.*, vol. 50, no. 14, pp. 994–996, Jul. 2014.

[8] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[9] J. Lee, M. Park, and C. Kim, *3D-CE1: Depth Intra Skip (DIS) Mode*, document JCT3V-K0033, Geneva, Switzerland, Feb. 2015.

[10] H. Liu and Y. Chen, "Generic segment-wise DC for 3D-HEVC depth intra coding," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 3219–3222.

[11] J. Y. Lee, H.-C. Wey, and D.-S. Park, "A fast and efficient multi-view depth image coding method based on temporal and inter-view correlations of texture images," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 12, pp. 1859–1868, Dec. 2011.

[12] L. Shen, Z. Zhang, and Z. Liu, "Inter mode selection for depth map coding in 3D video," *IEEE Trans. Consum. Electron.*, vol. 58, no. 3, pp. 1857–1865, Aug. 2012.

[13] L. Shen, P. An, Z. Liu, and Z. Zhang, "Low complexity depth coding assisted by coding information from color video," *IEEE Trans. Broadcast.*, vol. 60, no. 1, pp. 128–133, Mar. 2014.

[14] 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.

[15] Z. Pan, Y. Zhang, and S. Kwong, "Fast mode decision based on texture–depth correlation and motion prediction for multiview depth video coding," *J. Real Time Image Process.*, vol. 11, pp. 1–10, Mar. 2013, doi: 10.1007/s11554-013-0328-3.

[16] L. Shen and Z. Zhang, "Efficient depth coding in 3D video to minimize coding bitrate and complexity," *Multimedia Tools Appl.*, vol. 72, no. 2, pp. 1639–1652, Sep. 2014.

[17] 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.

[18] 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.

[19] 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, pp. 1–23, Feb. 2015.

[20] R. Conceicao, G. Avila, G. Correa, M. Porto, B. Zatt, and L. Agostini, "Complexity reduction for 3D-HEVC depth map coding based on early skip and early DIS scheme," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Phoenix, AZ, USA, Sep. 2016, pp. 1116–1120.

[21] J. Chen, B. Wang, H. Zeng, C. Cai, and K.-K. Ma, "Sum-of-gradient based fast intra coding in 3D-HEVC for depth map sequence (SOG-FDIC)," *J. Vis. Commun. Image Represent.*, vol. 48, pp. 329–339, May 2017.

[22] H.-B. Zhang, C.-H. Fu, Y.-L. Chan, S.-H. Tsang, and W.-C. Siu, "Probability-based depth intra-mode skipping strategy and novel VSO metric for DMM decision in 3D-HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 2, pp. 513–527, Feb. 2018.

[23] J. Lei, J. Duan, W. Feng, 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.

[24] M. Saldanha, G. Sanchez, C. Marcon, and L. Agostini, "Fast 3D-HEVC depth maps intra-frame prediction using data mining," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 1738–1742.

[25] L. Shen, K. Li, R. Feng, and Z. Liu, "Efficient intra mode selection for depth-map coding utilizing spatiotemporal, inter-component and inter-view correlations in 3D-HEVC," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4195–4206, Sep. 2018.

[26] H.-N. Wang and Q. Li, "Fast decision algorithm for intra mode in depth map of 3D-HEVC," in *Proc. IEEE 2nd Int. Conf. Inf. Commun. Signal Process. (ICICSP)*, Weihai, China, Sep. 2019, pp. 327–331.

[27] R. Zhang, K. Jia, P. Liu, and Z. Sun, "Edge-detection based fast intra-mode selection for depth map coding in 3D-HEVC," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Sydney, NSW, Australia, Dec. 2019, pp. 1–4.

[28] 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.

[29] Q. Zhang, Q. Yang, Y. Chang, W. Zhang, and Y. Gan, "Fast intra mode decision for depth coding in 3D-HEVC," *Multidimensional Syst. Signal Process.*, vol. 28, no. 4, pp. 1203–1226, Oct. 2017.

[30] 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, pp. 170–180, May 2017.

[31] 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.

[32] H.-S. Kim and R.-H. Park, "Fast CU partitioning algorithm for HEVC using an online-learning-based Bayesian decision rule," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 130–138, Jan. 2016.

[33] J. Zhang, S. Kwong, and X. Wang, "Two-stage fast inter CU decision for HEVC based on Bayesian method and conditional random fields," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 11, pp. 3223–3235, Nov. 2018.

[34] K. Goswami and B.-G. Kim, "A design of fast high-efficiency video coding scheme based on Markov chain Monte Carlo model and Bayesian classifier," *IEEE Trans. Ind. Electron.*, vol. 65, no. 11, pp. 8861–8871, Nov. 2018.

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

[36] *3D-HEVC Reference Software (HTM 16.1)*. Accessed: Oct. 2021. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_3DVCSoftware/tags/HTM-16.1/cfg/3D-HEVC/

[37] M. Tanimoto, T. Fujii, and K. Suzuki, *View Synthesis Algorithm in View Synthesis Reference Software 2.0 (VSRS 2.0)*, document ISO/IEC JTC1/SC29/WG11 document M16090, Lausanne, Switzerland, Feb. 2008.

[38] 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.

DONGYAO ZOU received the Ph.D. degree in circuits and systems from the Beijing University of Posts and Telecommunications, Beijing, China, in 2008. In 2011, he served as the Director for the Internet of Things Engineering Department, School of Computer and Communication Engineering, Zhengzhou University of Light Industry, where he has been engaged in teaching and research work, since July 2008. His research interests include 3D signal processing, machine learning, pattern recognition, video codec optimization, and multimedia communication.

PU DAI received the B.S. degree in electronic information engineering from Pingdingshan University, Pingdingshan, China, in 2020. He is currently pursuing the master's degree in signal and information processing with the School of Computer and Communication Engineering, Zhengzhou University of Light Industry, Zhengzhou. His current research interests include image processing, high efficiency video coding, versatile video coding, machine learning, and extensions of the versatile video coding.

QIUWEN ZHANG received the Ph.D. degree in communication and information systems from Shanghai University, Shanghai, China, in 2012. Since 2012, he has been with the Faculty of 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 papers 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.

● ● ●