

## Research Article

# A Computation Complexity Reduction of the Size Decision Algorithm in 3D-HEVC Depth Map Intracoding

Hamza Hamout <sup>1</sup> and Abderrahmane Elyousfi <sup>2</sup>

<sup>1</sup>Department of Computer Science, Higher School of Technology, Ibn-Zohr University, Guelmin 81000, Morocco

<sup>2</sup>Department of Computer Science, National Engineering School of Applied Sciences, Ibn-Zohr University, Agadir 80000, Morocco

Correspondence should be addressed to Hamza Hamout; hamzahamout@gmail.com

Received 9 August 2021; Revised 4 June 2022; Accepted 10 June 2022; Published 29 June 2022

Academic Editor: Patrick Seeling

Copyright © 2022 Hamza Hamout and Abderrahmane Elyousfi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new video format called a multiview video plus depth map is recently designed as the most efficient 3D video representation. 3D-high efficiency video coding (3D-HEVC) is the international standard for 3D video coding finalized in 2016. In 3D-HEVC intracoding, the depth map is an essential component, in which the depth map intraprediction occupies more than 85% of the overall intraencoding time. In 3D-HEVC design, a highly flexible quadtree coding unit partitioning is adopted, in which one coding tree unit is partitioned recursively into a prediction unit (PU) from  $64 \times 64$  down to  $4 \times 4$ . This highly flexible partitioning provides more accurate prediction signals and thereby achieves better intradepth map compression efficiency. However, performing all depth map intraprediction modes for each PU level achieves high depth map intracoding efficiency, but it results in a substantial computational complexity increase. This paper proposes an amelioration of the previously proposed depth map PU size decision using an efficient homogeneity determination. The resulting experiences show that the proposed method can significantly save the computational complexity with a negligible loss of intracoding efficiency.

## 1. Introduction

In recent years, the digital 3D video and electronic device with a high resolution receive too much attention from the consumer and electronic applications [1]. To support the explosive increase in 3D video service, transmission, and large storage, the joint collaborative team on 3D video coding (JCT-3V) [2] developed the emerging international standard for 3D video coding, namely, 3D-high efficiency video coding (3D-HEVC) [3]. For displaying the 3D video on an autostereoscopic display and reducing the rate data significantly, a new alternative 3D video format for 3D scene representation is adopted by 3D-HEVC, commonly known as multiview video plus depth (MVD) maps [4, 5]. In an MVD format, only a few 2D texture videos and their corresponding depth maps are multiplexed into 3D video bitstream. At the decoded side, the additional intermediate views suitable for enabling the autostereoscopic view can be easily created by the synthesizing process using depth image-based rendering (DIBR) [6].

In the 3D video coding standard, 3D-HEVC uses a highly flexible coding unit recursive partition called a coding tree unit (CTU) [7], where each CTU is further split to create a quadtree structure based on new unit types: coding unit and prediction unit [8]. It is allowed to encode more efficiency in all depth map characteristics and then reduce the bandwidth. This highly flexible coding unit leads to a vast coding complexity increase caused by the extensive search for the best size using rate-distortion optimization cost [9]. Hence, it is imperative to propose an efficient and fast 3D video coding algorithm to decrease this coding complexity.

Recently, some research studies have been developed to accelerate the size decision process in 3D-HEVC intradepth map coding [10–17] and in HEVC/VVC [18, 19]. In [10], the proposed method used a specific feature called a corner point to accelerate the quadtree intradecision. In [11], the authors propose a strategy that speeds up the depth map intraprediction by ignoring the small CU size using a threshold variance value. Chen et al. [12] develop an efficient method based on the gradient's sum to decide whether the

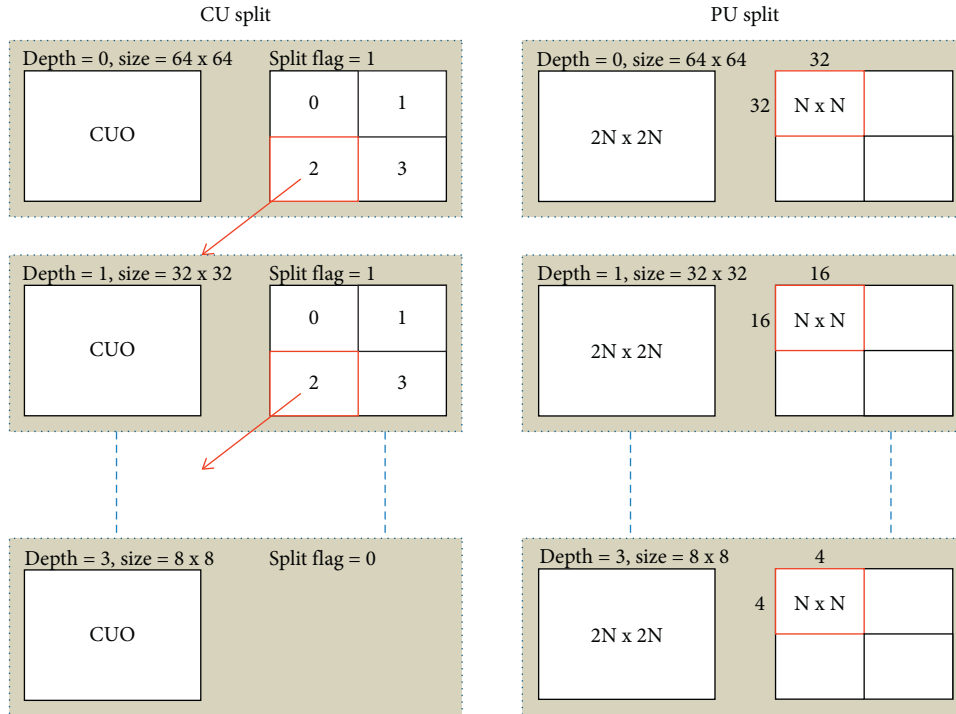


FIGURE 1: The architecture of the recursive split of the coding unit and the prediction unit.

CU must skip unnecessary rate-distortion optimization cost of the smaller partition sizes. The proposed algorithm in [13] exploits the correlation between hierarchical CU/PU and then reduces the number of recursive rate-distortion optimization costs in the depth map intraprediction. Chiang et al. [14] develop a low complexity CU size prediction. The smaller CU sizes are not performed if some conditions are satisfied, including the oversplitting of the collocated texture CU. The authors in [15] exploit the statistical data analysis to create fast CU/PU size decisions to reduce the complexity of depth map intracoding. In our previously proposed algorithm [16], the depth map intrasize decision is established, including thresholds calculated based on an efficient clustering algorithm and tensor features. In [17], a low complexity intramode selection algorithm is proposed to reduce complexity of depth intraprediction in both intraframes and interframes. In [18], a novel fast QTMT partition decision framework is developed, which can determine the partition decision on the both QT and multitype tree with a novel cascade decision structure. In [19], the authors propose a fast algorithm for VVC from two aspects of mode selection and prediction terminating to reduce coding complexity.

This paper proposed an efficient size intramodel decision as an amelioration of our previous work [16] using an efficient clustering algorithm, namely, automatic merging possibilistic clustering method [20] and new features. The proposed method decreases the synthesized views efficiency losses while conserving the same coding time savings. For this goal, we preserve the same previous proposed algorithm with different features.

The remainder of this paper is as follows: Section 2 presents an overview of the size coding tools in 3D-HEVC depth map intracoding. Section 3 gives a brief introduction of the AM-

PCM and feature extraction. Then, the statistical analysis and the proposed algorithm are discussed in Section 4. The resulting experiences are presented and discussed in Section 5. The conclusion of this paper is presented in Section 6.

## 2. Overview of 3D-HEVC Depth Map Intraprediction

In this section, we give a description of the highly flexible coding structure followed by the intraprediction modes in 3D-HEVC intracoding.

*2.1. Highly Flexible Coding Units.* To achieve the highest 3D-HEVC depth map intraprediction efficiency, the 3D-HEVC video encoder exploits spatial correlation inside the depth map views based on highly flexible coding structure units, which are described as follows: each depth map view is partitioned into a coding tree unit (CTU) [21], and each CTU is performed using three basic units, namely, coding unit (CU), prediction unit (PU), and transform unit (TU). Each one of these units has a different surly role. The CU is the basic coding unit with a square size from  $64 \times 64$  recursively subdivided down to  $8 \times 8$ . It consists of one luma and two chromas [22], and it is further split into PUs (see Figure 1) and TUs.

The prediction unit is defined by a subdivision of the coding unit where the same intraprediction mode is applied [23]. The PU is used for carrying the information associated with the intraprediction process, in which each PU supports two different splits,  $2N \times 2N$  corresponds to  $64 \times 64$ ,  $32 \times 32$ ,  $16 \times 16$ , and  $8 \times 8$ , and  $N \times N$  corresponds to  $4 \times 4$  [24], as depicted in Figure 1.

TABLE 1: Size of the blocks for different unit types.

| Size of blocks | Depth level | CU | PU | TU |
|----------------|-------------|----|----|----|
| $64 \times 64$ | 0           | ✓  | ✓  | —  |
| $32 \times 32$ | 1           | ✓  | ✓  | ✓  |
| $16 \times 16$ | 2           | ✓  | ✓  | ✓  |
| $8 \times 8$   | 3           | ✓  | ✓  | ✓  |
| $4 \times 4$   | 4           | —  | ✓  | ✓  |

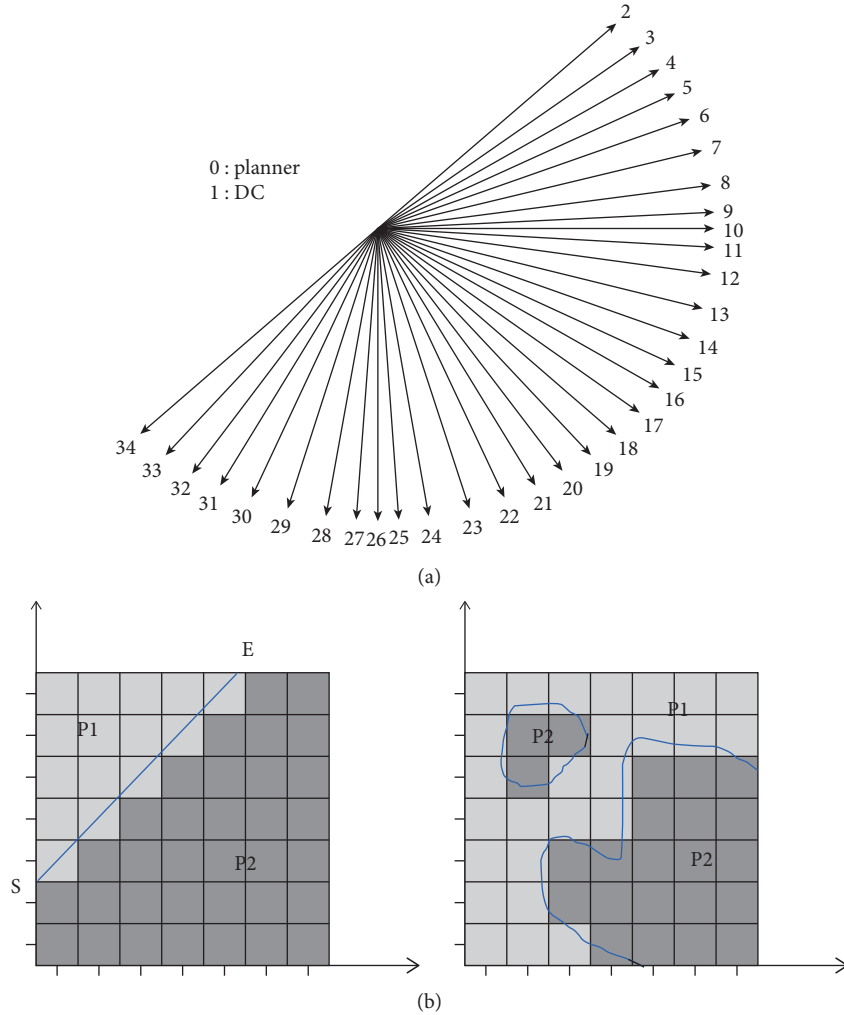


FIGURE 2: 3D-HEVC depth map intraprediction tools: (a) intra-angular and (b) depth modeling modes: wedgelet (left) and contour (right).

The transform unit is formed by the difference in pixels between block Original and block Prediction. It shares the same procedures of the rate-distortion optimization cost. Each transform unit allows the sizes from a maximum of  $32 \times 32$  to a minimum of  $4 \times 4$ . The multiple TU split forms another quadtree called residual quadtree (RQT). Table 1 presents the supported sizes for the CU, PU, and TU, respectively.

**2.2. 3D-HEVC Depth Map Intraprediction Tools.** The 3D-HEVC depth map intraprediction exploits the previously coded block to remove the spatial redundancies [25, 26]

where each depth map PU is evaluated by rough mode decision [27]. In the rough mode decision, 35 modes including DC/Planar and 33 angular directions (see Figure 2 (top)) are performed based on the sum of absolute transformed difference (SATD) [28]. The only subset of the modes having the lowest SATDs is selected to be added into a full rate-distortion list. Next, most probable modes are performed and three intramodes derived from the left and top of already coded neighboring PUs are added to the full rate-distortion. Finally, the depth modeling mode is accomplished by approximating a model pattern that partitions the depth map PU into two nonrectangular regions  $P_1$  and  $P_2$  according to two partition types: wedgelet partition

(DMM<sub>1</sub>) and contour partition (DMM<sub>4</sub>) as shown in Figure 2 (bottom). The result of depth modeling modes ((DMM<sub>1</sub>) and (DMM<sub>2</sub>)) is added to the full rate-distortion list only if some conditions are satisfied [29]. Figure 3 illustrates the depth map intraprediction process in 3D-HEVC.

In 3D-HEVC, a single coding tree unit comprises 341 PUs, in which the sizes vary from 64 × 64 down to 4 × 4. For each PU, rough mode decision, most probable modes, and depth modeling modes are evaluated to select the best depth map intraprediction mode. Table 2 presents the detailed depth map intraprediction modes for a single coding tree unit.

### 3. Clustering Algorithm and Feature Extraction

Clustering is a method that regroups the data with similar characteristics in the same cluster. In the fuzzy clustering algorithms, in [20] an efficient clustering algorithm is proposed, namely, automatic merging possibilistic clustering method (AM-PCM). The AM-PCM overcame the weakness of the previous fuzzy clustering systems like the fuzzy C-mean [30] and possibilistic C-mean [31], and it is very robust in a noisy environment and resolves the parameter and initialization problems.

The purpose of this work is to decrease the high complexity selection of the best size in the depth map intraprediction while preserving the rate-distortion performance. In the CTU, the selection of the best size is mostly related to the depth map content. This relationship is well illustrated in Figure 4, in which the homogeneous CTU is often coded with large sizes and the complex CTU is coded with small sizes. However, computing the features that can describe the homogeneous depth map CTU can be used to predict the suitable sizes to be evaluated. In this paper, we use features different from those used in our previous contribution [16]. Hence, we select the amplitude of simplified mass center (ASMCV) and the variance [32] as the proposed homogeneous features. Let  $L(i, j)$  be the luminance value of the pixel at  $(i, j)$  and the CTU, the ASMCV, and the variance are, respectively, defined by equations (1) and (2).

ASMCV =  $D_x + D_y$ , where

$$D_x = \sum_{j=0}^{31} \sum_{i=0}^{63} \left( \frac{64}{2} - j \right) \times |L(i, j) - L(i, 64 - j - 1)|, \quad (1)$$

$$D_y = \sum_{i=0}^{31} \sum_{j=0}^{63} \left( \frac{64}{2} - i \right) \times |L(i, j) - L(64 - i - 1, j)|,$$

$$\text{Var} = \frac{1}{64 \times 64} \sum_{i=0}^{63} \sum_{j=0}^{63} [L(i, j)]^2 - \left[ \frac{1}{64 \times 64} \sum_{i=0}^{63} \sum_{j=0}^{63} L(i, j) \right]^2. \quad (2)$$

Equation (3) presents the amplitude of the feature vector (ASMCV and Var). Theoretically, the amplitude of the computed features is correlated to the CTU sizes prediction in the 3D-HEVC depth map intraprediction, where if the

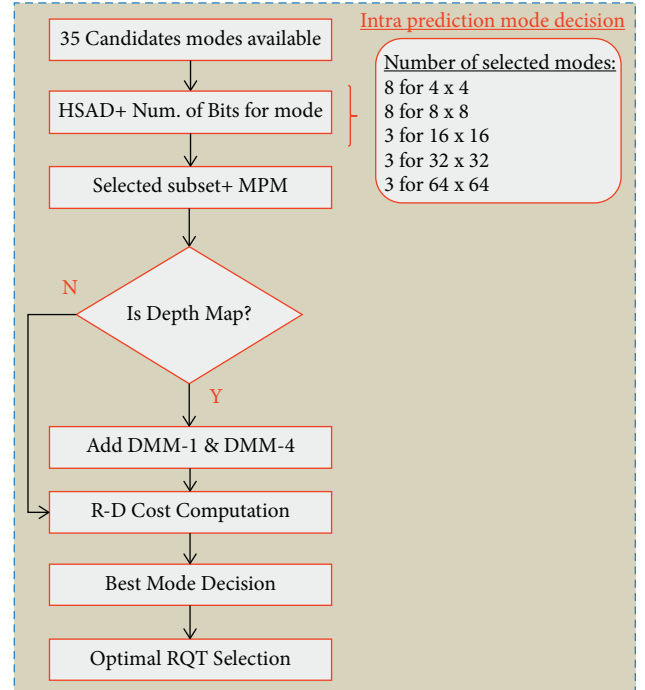


FIGURE 3: The depth map intraprediction process in 3D-HEVC.

TABLE 2: Depth map intraprediction mode analysis.

| Size | Number PU | Intra-angular | 3D-HEVC |        | Total RDO-cost |
|------|-----------|---------------|---------|--------|----------------|
|      |           |               | MPMs    | DMMs   |                |
| 64   | 1         | 3             | 3       | 0      | 6              |
| 32   | 4         | 3             | 3       | 0 or 2 | 24–32          |
| 16   | 16        | 3             | 3       | 0 or 2 | 96–128         |
| 8    | 64        | 8             | 3       | 0 or 2 | 704–832        |
| 4    | 256       | 8             | 3       | 0 or 2 | 2816–3328      |

amplitude is high, then the small sizes are likely to be selected. On another side, if the amplitude is low, then the larger sizes are likely to be selected. Figure 5 presents the histogram distribution size of the depth map CTU selected by the anchor 3D-HEVC according to the amplitude of the homogeneous features. In Figure 5, the case where the amplitude is distributed between 0 and 1000 (low amplitude), more than 97% of the CTU are coded with 64 × 64 and the probability to select the other sizes are almost insignificant. In another case, where the amplitude increases, the probability that the CTU will be coded with small sizes increases and the probability of those with large sizes decreases, which justifies our theory.

$$\text{Amp} = \sqrt{\text{ASMCV}^2 + \text{Var}^2}. \quad (3)$$

### 4. Clustering Data Analysis and Proposed Size Prediction in Depth Map Intra-coding

For an efficient data clustering result, the clustering method input data must consider all the possible contents of the

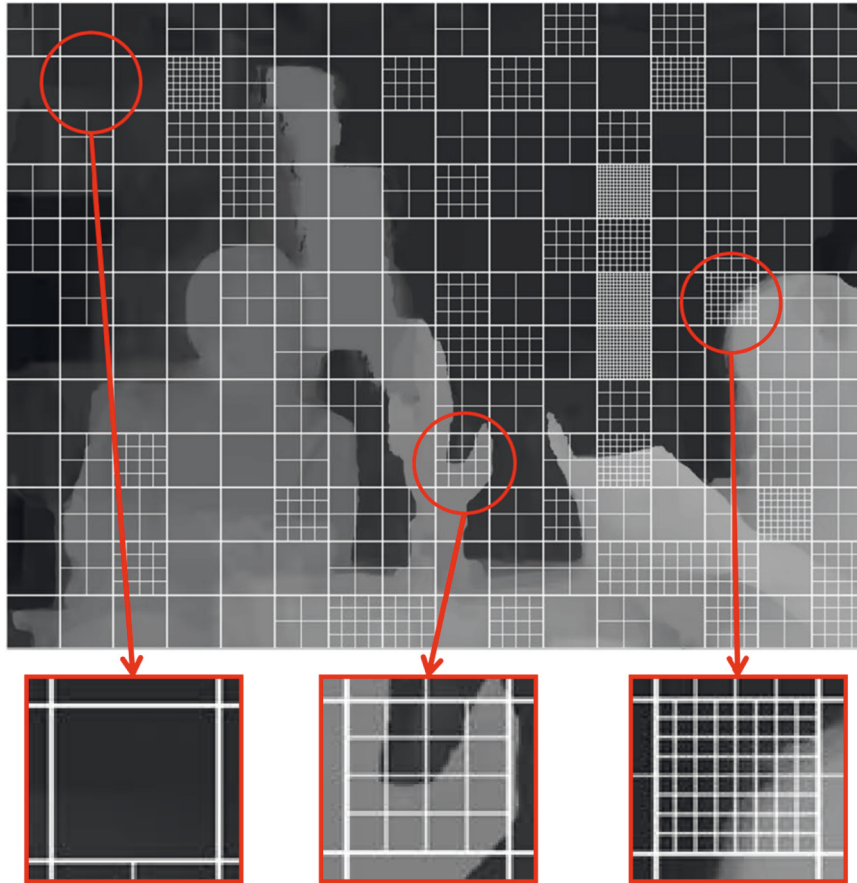


FIGURE 4: The relationship between the CTU and its complexity content.

depth map. The input data for the AM-PCM were regrouped from a set of depth map sequences, which are “LoveBrid1” with resolution  $1024 \times 768$ , “Akko&Kayo” and “Rena” with resolution  $640 \times 480$  and “Champagne\_tower,” “Dog,” and “Pantomime” with resolution  $1280 \times 960$ . The primary dataset contains CTUs of the first frame from every 24th frames of all depth map sequences, that is, a total of 53,521 CTUs. Taking into consideration all possible homogeneous and complex depth map contents, we compute for all 53,521 CTUs in ASMCV equation (1) and variance equation (2) and sorted them according to equation (3). Next, we select the input AM-PCM clustering data as the feature vectors by a frequency of 13, a total of 4117 vectors, which present only 7.7% of the primary dataset. In the output of the AM-PCM, we get exactly 1117 cluster centers sorted according to their amplitude.

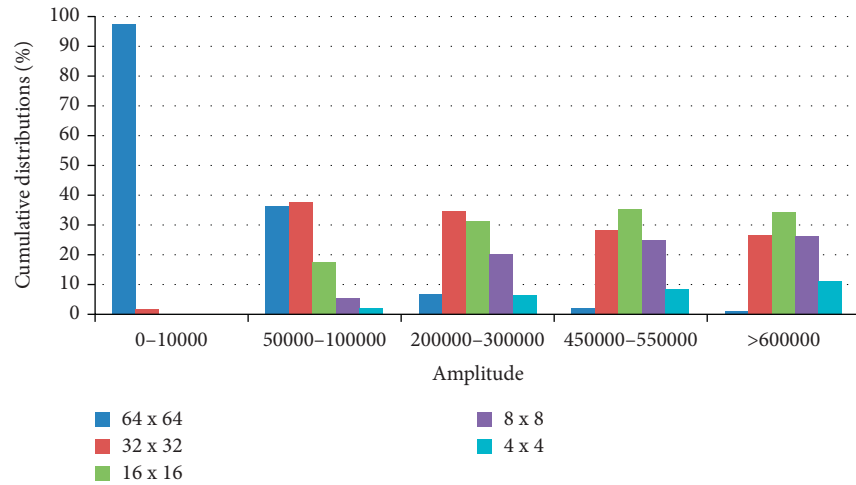
The purpose of this article is to create a size decision model whose goal is to estimate the best sizes for a given CTU of the depth map and then decrease the 3D-HEVC depth map intraprediction complexity. Using the CTU of the six sequences, a total of 1,254,240 vectors, where the ASMCV and variance compose each vector and the size flag, we start combining all 1,254,240 vectors and the output of the AMP-PCM (1117 cluster centers) by distributing each vector to its cluster center that minimizes the Euclidean distance. According to the percentage distribution of the CTU sizes in each cluster, we determine which sizes are dominant. If the

percentage is more significant than 10%, it will be considered, or else it will be ignored. Regrouping the clusters that have a similar dominant size, we get five groups. This process is executed for all depth map quantization parameters. Table 3 presents the results of this distribution process.

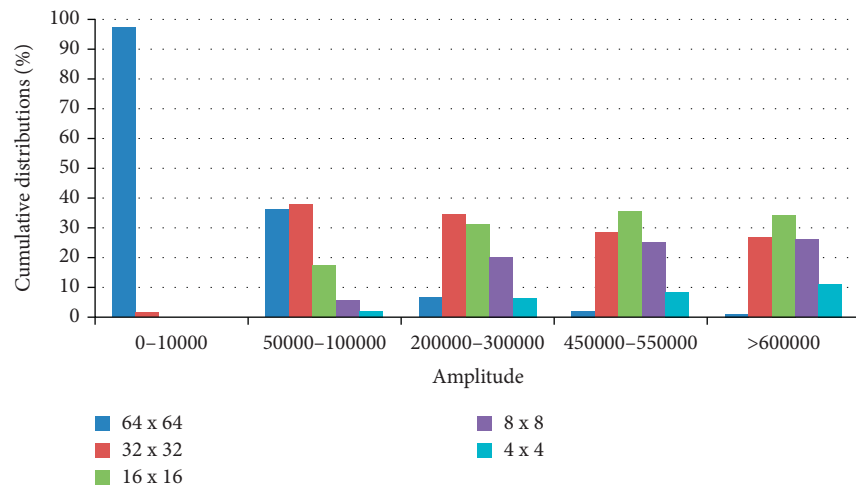
Using the results presented in Table 3, a size decision model for depth map intraprediction in 3D-HEVC is defined based on the dominant size distributions. Hence, Table 4 presents the size prediction model according to four thresholds  $Th_i$ ,  $i \in \{1, 2, 3, 4\}$ . The thresholds are determined using the amplitude of the regrouped center clusters, where they are computed as follows: as mentioned previously, the center clusters are sorted according to their amplitude (equation (3)), so let be  $R_i$ ,  $i \in \{1, 2, 3, 4, 5\}$  the regrouping clusters centers, we set  $Th_i = (\text{MinAmp}(R_{i+1}) + (\text{MaxAmp}(R_i)))/2$ , where  $\text{MinAmp}$  and  $\text{MaxAmp}$  return, respectively, the minimum and the maximum amplitude of the center clusters regrouped in a specific  $R_i$ . Table 5 presents the respective threshold values used in the proposed size model decision for depth map intracoding according to the depth map quantization parameter. The algorithm of the proposed size model decision is presented in algorithm 1.

## 5. Resulting Experiences

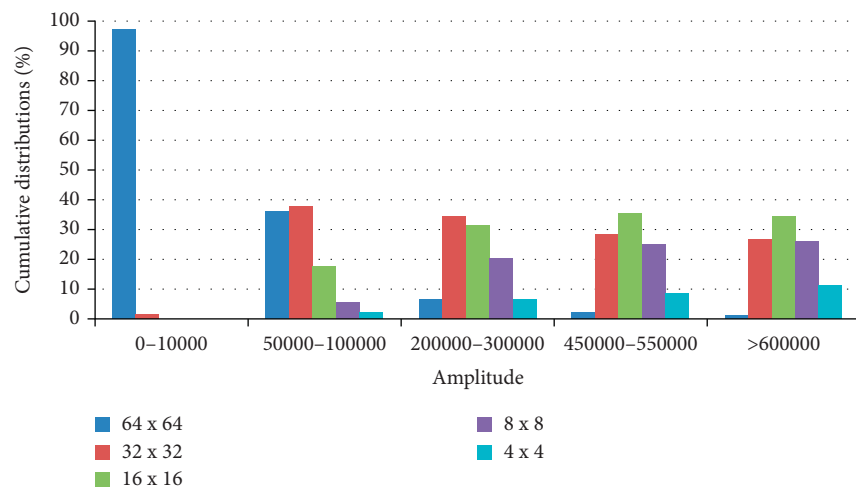
This section presents the resulting experiences of the size decision model performance developed for the depth map



(a)



(b)



(c)

FIGURE 5: The histogram distribution size of the depth map CTU according to the amplitude and  $QP_{\text{DepthMap}}$  34, 39, and 42. (a)  $QP_{\text{DepthMap}} = 34$ . (b)  $QP_{\text{DepthMap}} = 39$ . (c)  $QP_{\text{DepthMap}} = 42$ .

intraprediction in 3D-HEVC compared with the state-of-the-art size prediction methods [10–16]. The anchor HTM-16.2 is used as the reference 3D video coding software [33].

The experimental configurations are based on JCT3V-G1100 documents [34], where the experiments are conducted on eight sequences “Shark,” “UndoDancer,” “PoznanStreet,”

TABLE 3: The results of the distribution process for all depth map QPs.

| Region | QP | Depth map prediction unit size |             |             |           |           |
|--------|----|--------------------------------|-------------|-------------|-----------|-----------|
|        |    | 64 × 64 (%)                    | 32 × 32 (%) | 16 × 16 (%) | 8 × 8 (%) | 4 × 4 (%) |
| 1      | 34 | 98.49                          | 1.31        | 0.11        | 0.08      | 0.00      |
|        | 39 | 98.40                          | 1.43        | 0.12        | 0.04      | 0.01      |
|        | 42 | 97.31                          | 2.14        | 0.43        | 0.11      | 0.01      |
|        | 45 | 96.41                          | 3.19        | 0.26        | 0.12      | 0.02      |
| 2      | 34 | 63.13                          | 25.43       | 7.47        | 2.76      | 1.21      |
|        | 39 | 68.11                          | 20.97       | 7.63        | 2.62      | 0.67      |
|        | 42 | 66.92                          | 24.34       | 6.11        | 2.16      | 0.47      |
|        | 45 | 72.14                          | 20.75       | 4.84        | 2.01      | 0.26      |
| 3      | 34 | 3.98                           | 39.54       | 17.61       | 6.58      | 2.29      |
|        | 39 | 28.96                          | 41.24       | 19.79       | 7.92      | 2.09      |
|        | 42 | 32.32                          | 43.63       | 16.41       | 6.17      | 1.47      |
|        | 45 | 30.04                          | 47.73       | 16.57       | 5.04      | 0.63      |
| 4      | 34 | 14.72                          | 45.27       | 24.48       | 12.33     | 3.21      |
|        | 39 | 10.14                          | 47.93       | 24.54       | 13.94     | 3.44      |
|        | 42 | 10.26                          | 51.08       | 24.85       | 10.16     | 3.65      |
|        | 45 | 18.02                          | 47.99       | 24.35       | 8.21      | 1.42      |
| 5      | 34 | 3.70                           | 30.37       | 33.50       | 23.40     | 9.03      |
|        | 39 | 4.88                           | 40.94       | 31.99       | 17.01     | 5.17      |
|        | 42 | 8.16                           | 44.13       | 30.87       | 13.89     | 2.96      |
|        | 45 | 2.03                           | 61.11       | 29.01       | 7.22      | 0.63      |

TABLE 4: Depth map size decision model.

| Amplitude                               | Size in intraprediction |         |         |       |       |
|---|-------------------------|---------|---------|-------|-------|
|   | 64 × 64                 | 32 × 32 | 16 × 16 | 8 × 8 | 4 × 4 |
| Amp < Th <sub>1</sub>                   | ✓                       | —       | —       | —     | —     |
| Th <sub>1</sub> ≤ Amp < Th <sub>2</sub> | ✓                       | ✓       | —       | —     | —     |
| Th <sub>2</sub> ≤ Amp < Th <sub>3</sub> | ✓                       | ✓       | ✓       | —     | —     |
| Th <sub>3</sub> ≤ Amp < Th <sub>4</sub> | ✓                       | ✓       | ✓       | ✓     | —     |
| Amp ≥ Th <sub>4</sub>                   | —                       | ✓       | ✓       | ✓     | ✓     |

TABLE 5: The size decision model thresholds for each depth map quantization parameters.

| QP | Thresholds      |                 |                 |                 |
|----|-----------------|-----------------|-----------------|-----------------|
|    | Th <sub>1</sub> | Th <sub>2</sub> | Th <sub>3</sub> | Th <sub>4</sub> |
| 34 | 2853.65         | 28624.94        | 142264.87       | 170287.95       |
| 39 | 15342.66        | 154886.01       | 481906.59       | 535758.73       |
| 42 | 70421.92        | 377178.31       | 1134863.28      | 1318718.60      |
| 45 | 159877.88       | 942726.90       | 2628393.57      | 4235644.50      |

“PoznanHall2,” “GTFly,” “Newspaper1,” “Kendo,” and “Balloons” in all intraconfiguration using three textures and their associated depth map. The quantization parameters are (25, 34), (30, 39), (35, 42), and (40, 45) for the texture and depth map. The depth map efficiency is performed using Bjontegaard metrics [35] and time saving.

Table 6 presents the resulting experiences of the proposed size decision model and the state of the art [10–16], where BDBR shows the Bjontegaard bit-rate performance using the Y-PSNR of the six synthesized views over the total bit-rate. The time reduction (TR) presents the time saving achieved by the state of the art and the proposed size decision model. From Table 6, the proposed method can decrease the computational complexity for eight sequences

compared to the reference software. The time saving is often high for sequences with low motion like “PoznanHall2” (51.6%) and low for complex sequences like “Newspaper1” (32.3%). In conclusion, the proposed size decision model skips unnecessary size partition evaluation for each CTU.

Figure 6 illustrates the time saving and the rate-distortion curves of the proposed size decision model in all intraconfiguration in comparison with the reference software HTM-16.2 for three sequences “GTFly,” “Balloons,” and “Newspaper1.” As presented in Figure 6, the proposed size decision model rate-distortion curve is analogous to the anchor HTM-16.2 from high to low bit-rate; the proposed method improves the time significantly, saving for all quantization parameters.

```

(1) Input: coding tree unit
(2) Compute the ASMCV of the CTU by equation (2);
(3) Compute the variance of the CTU by equation (1);
(4) Compute the amplitude by equation (3);
(5) If Amp < Th1then
(6) Evaluate the rate-distortion cost only for 64 × 64;
(7) else if Th1 ≤ Amp < Th2then
(8) Evaluate the rate-distortion cost only for 64 × 64 and 32 × 32;
(9) else if Th2 ≤ Amp < Th3then
(10) Evaluate the rate-distortion cost only for 64 × 64, 32 × 32 and 16 × 16;
(11) else if Th3 ≤ Amp < Th4then
(12) Evaluate the rate-distortion cost only for 64 × 64, 32 × 32, 16 × 16 and 8 × 8;
(13) else if Amp ≥ Th4then
(14) Evaluate the rate-distortion cost only for 32 × 32, 16 × 16, 8 × 8 and 4 × 4;
(15) end
(16) end
(17) end
(18) end
(19) end

```

ALGORITHM 1: Proposed size model decision for 3D-HEVC depth map intra coding.

TABLE 6: Resulting experiences of the proposed size decision model in all intraconfiguration.

| Sequence     | PM1: [10] |        | PM2: [11] |        | PM3: [12] |        | PM4: [13] |        | PM5: [14] |        | PM6: [15] |        | PM7: [16] |        | Proposed |        |
|--------------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|----------|--------|
|              | BDBR (%)  | TR (%) | BDBR (%)  | TR (%) | BDBR (%)  | TR (%) | BDBR (%)  | TR (%) | BDBR (%)  | TR (%) | BDBR (%)  | TR (%) | BDBR (%)  | TR (%) | BDBR (%) | TR (%) |
| Balloons     | 0.29      | 41.0   | -0.10     | 24.4   | 0.00      | 25.8   | 0.05      | 20.0   | 0.05      | 23.6   | X         | X      | 0.35      | 37.6   | 0.12     | 32.9   |
| Kendo        | 0.29      | 39.0   | 0.10      | 22.7   | 0.10      | 28.0   | 0.12      | 22.6   | 0.03      | 26.3   | 0.71      | 47.1   | 0.43      | 38.7   | 0.17     | 35.2   |
| Newspaper1   | -0.23     | 36.0   | 0.20      | 21.1   | 0.20      | 22.3   | 0.10      | 11.1   | 0.21      | 22.8   | 1.61      | 47.9   | 0.43      | 43.5   | 0.08     | 32.3   |
| GTFly        | 0.20      | 45.0   | 0.20      | 40.2   | 0.00      | 13.0   | 0.10      | 20.7   | 0.03      | 34.4   | 0.32      | 47.9   | 0.32      | 49.6   | 0.08     | 35.0   |
| Poznanhall2  | 0.33      | 48.0   | 0.60      | 39.9   | 0.10      | 34.1   | 0.40      | 45.3   | 0.67      | 46.7   | X         | X      | 0.51      | 54.0   | 0.39     | 51.6   |
| Poznanstreet | 1.16      | 40.0   | 0.80      | 32.4   | 0.00      | 15.0   | 0.28      | 17.9   | 0.12      | 29.5   | 0.53      | 50.5   | 0.55      | 49.0   | 0.26     | 41.6   |
| Undodancer   | 1.01      | 39.0   | 0.40      | 44.1   | 0.00      | 17.6   | 0.27      | 36.1   | 0.07      | 38.0   | 1.02      | 45.5   | 0.53      | 58.3   | 0.29     | 49.3   |
| Shark        | X         | X      | 0.10      | 17.0   | 0.00      | 19.0   | 0.13      | 31.3   | 0.07      | 32.1   | X         | X      | 0.41      | 39.7   | 0.26     | 44.0   |
| 1024 × 768   | 0.12      | 39.0   | 0.07      | 22.7   | 0.10      | 25.4   | 0.09      | 17.9   | 0.10      | 24.2   | 1.016     | 47.5   | 0.40      | 39.9   | 0.12     | 33.5   |
| 1920 × 1088  | 0.68      | 43.0   | 0.42      | 34.7   | 0.02      | 19.7   | 0.24      | 30.3   | 0.19      | 36.1   | 0.62      | 47.9   | 0.47      | 58.1   | 0.26     | 51.0   |
| Average      | 0.44      | 41.0   | 0.29      | 30.2   | 0.05      | 21.9   | 0.18      | 25.6   | 0.16      | 31.7   | 0.84      | 47.8   | 0.44      | 46.3   | 0.21     | 40.2   |

To perform more on our proposed method and compare it with the state-of-the-art size prediction methods, we visualize the BDBR-TR as a two-dimensional figure, in which we present in Figure 7, the TR at the cost of BDBR for the proposed method and the state-of-the-art size prediction methods for four sequences “Kendo,” “GTFly,” “Poznan-Street,” and “UndoDancer.” As illustrated in Figure 7, it is clear that our proposed method has a very good balance between time saving and BDBR compared to all state-of-the-art methods, which makes our proposed method more efficient.

Evaluating the subjective visual quality, Figure 8 presents the renderer views for “UndoDancer” and “PoznanStreet” sequences utilizing the proposed algorithm in comparison with the renderer views using the anchor HTM-16.2. From Figure 8, we observed that the proposed size decision model preserves almost the same synthesized view quality visual compared to the 3D video coding reference software, especially around the object boundaries where an occlusion

zone frequently arises in the synthesized views. The rest of the sequences present the same subjective visual quality.

Comparing the proposed size decision model developed for depth map intraprediction with the state-of-the-art size prediction methods that also develop solutions to decrease the size prediction in the depth map intraprediction [10–16], the proposed size decision model reaches a significant encoded time reduction with a negligible BDBR increase.

Comparing our proposed method to the state-of-the-art size prediction methods, our algorithm achieves almost the same time saving to the one presented in [10], in which the technique developed in [10] yields 41.0% in the time saving and a 0.44% increase in BDBR in comparison with HTM-12.1 as the reference software. In contrast, our method is implemented in HTM-16.2, including all 3D coding tools from HTM-12.1.

In [11], the proposed algorithm reaches 30.2% in the encoding time saving with a 0.29% increase in BDBR compared with HTM-13.0. Thus, the algorithm presented in



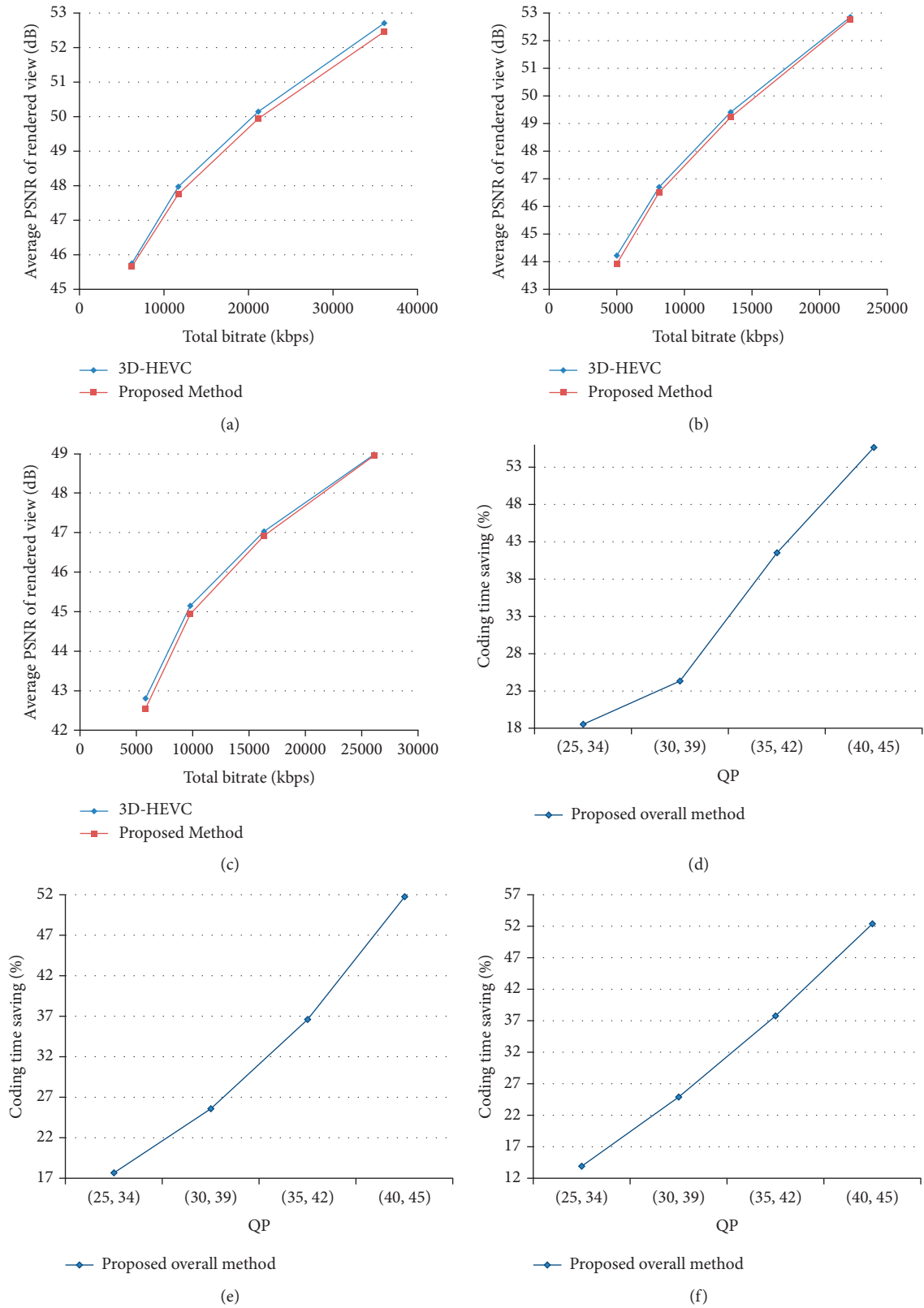


FIGURE 6: Rate-distortion and time-saving curves of “GTFly,” “Balloons,” and “Newspaper1” under different QPs in all intraconfiguration. (a) Rate-distortion curve of GTFly. (b) Rate-distortion curve of Balloons. (c) Rate-distortion curve of Newspaper1. (d) Time reduction curve of GTFly. (e) Time reduction curve of Balloons. (f) Time reduction curve of Newspaper1.

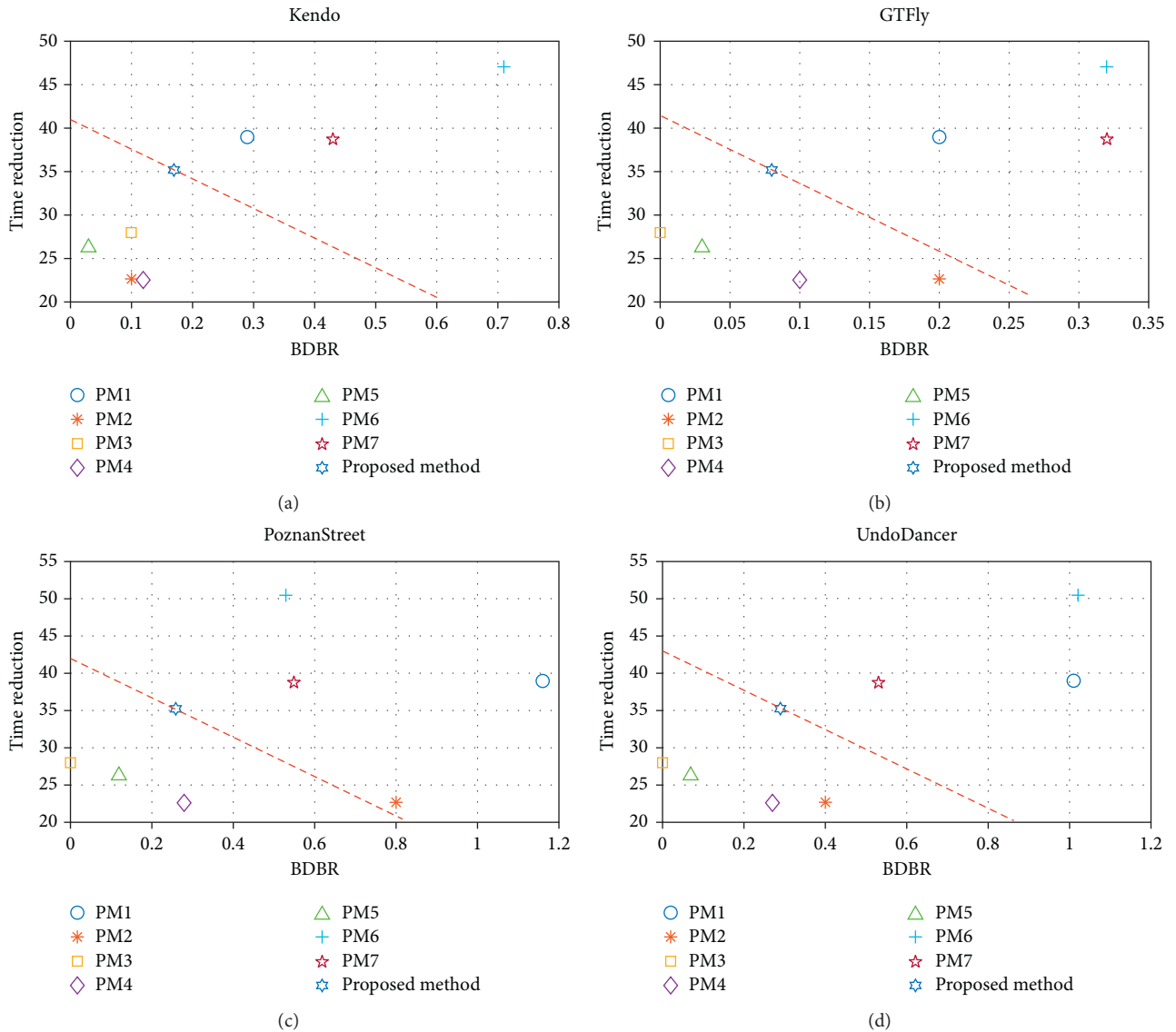


FIGURE 7: BDBR in function for time reduction of Kendo, GTFly, PoznanStreet, and UndoDancer.

[11] evaluates only 100 frames for each sequence. Though our algorithm is evaluated in full frames for each sequence, it is not feasible to make a rational comparison with the proposed method. To overcome this issue, additional experiments on the first 100 frames using our proposed algorithm are conducted. Our proposed method achieves 38.4% in the encoding time saving with a 0.19% increase in BDBR compared with HTM-16.2. Comparing our proposed method to the algorithm proposed in [11], our proposed algorithm surpasses the method proposed in [11] by a factor of 1.3 and 0.6, respectively, for the encoding time saving and the BDBR.

Concerning the algorithm developed in [12], the method proposed in this paper reduces the computational complexity significantly than that reported in [12] by 21.9%. Regarding the encoding performance, the technique in [12] preserves almost the same reconstructed synthesized view efficiency compared with HTM-16.1.

Concerning the method developed by Kim et al. in [13], the proposed algorithm reaches a lower encoding time reduction for all eight sequences in comparison with HTM-12.0. Compared with the proposed size decision model, our algorithm achieves a higher time reduction by a factor of 1.6. On the other hand, the proposed method in [13] surpasses our algorithm in the BDBR where it increased by 0.18%.

Taking into consideration the method in [14], the algorithm in this paper surpasses the results presented in [14] in terms of time reduction by a factor of 1.3, where the method in [14] achieves 31.7% in time reduction and a 0.16% increase in BDBR when compared to HTM-15.0.

Compared to the method developed by Gang et al. in [15], our proposed size decision model reaches almost the same time reduction as [15], which achieves 47.8% in time reduction in comparison with HTM-13.0. Taking into consideration the coding efficiency, the work in [15] presents a 0.84% increase in BDBR when compared to HTM-13.0.

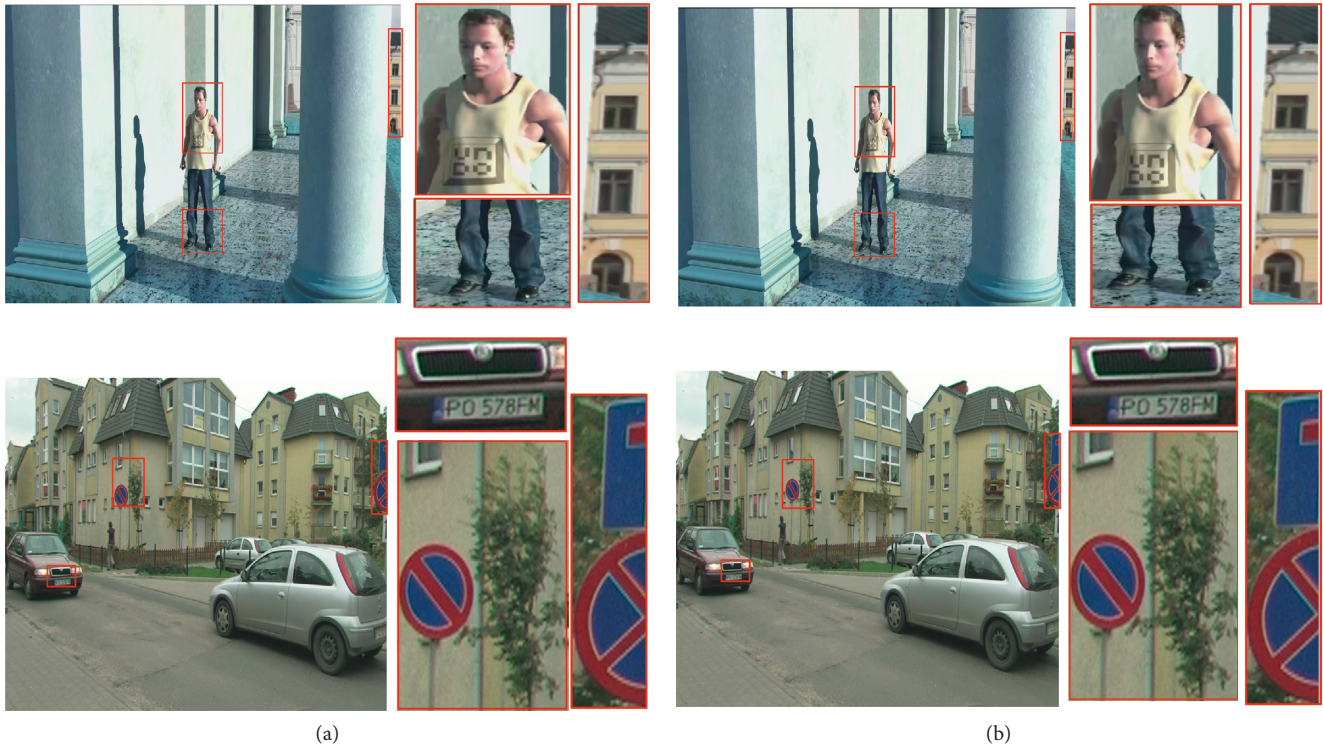


FIGURE 8: The comparison of the subjective visual quality of the synthesized views of the proposed method (b) and the reference software (a) for UndoDancer and PonznanStreet.

Since the method in [15] does not give all eight sequences performance, there is no fear of making a reasonable comparison with the present method.

Taking into consideration our previous size decision model developed in [16], the size decision model proposed in this paper reaches almost the same time reduction as [16] by a factor of 0.9. Regarding the coding efficiency BDBR, the proposed size decision presented in this paper outperforms our previous size decision model by a factor of 0.5. Our previous size decision model results in a 0.44% increase, which is high compared to the current size decision model.

## 6. Conclusion

In this paper, a computation complexity reduction of the size decision algorithm for the depth map intraprediction in 3D-HEVC has been developed. In brief, the proposed approach applies an efficient clustering algorithm, namely, automatic merging possibilistic clustering method based on a set of selective data. Using the clustering algorithm and data analysis output, we create our proposed size decision model for the depth map intraprediction. Finally, this model predicts the size flags and then accelerates the 3D-HEVC depth map intracoding.

Taking the result experience, the proposed size decision model achieves 40.2% in time reduction with a negligible increase in BDBR by 0.21% in comparison with the anchor HTM-16.2. Compared to the state-of-the-art size prediction methods, the results presented in this paper confirm that the proposed size decision model can significantly reduce the

depth map coding time while preserving almost the same rate-distortion as the original 3D video coding standard.

## Data Availability

The data used in this study are described in the following reference: “K. Muller, A. Vetro, Joint Collaborative Team on 3D Video Coding Extension Development Document JCT3V-G1100, 7th Meeting, San Jose, US (2014).”

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] L. Shen, Z. Zhang, and P. An, “Fast cu size decision and mode decision algorithm for hevc intra coding,” *IEEE Transactions on Consumer Electronics*, vol. 59, no. 1, pp. 207–213, 2013.
- [2] 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 Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 35–49, 2016.
- [3] K. Müller, H. Schwarz, D. Marpe et al., “3d high-efficiency video coding for multi-view video and depth data,” *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3366–3378, 2013.
- [4] P. Kauff, N. Atzpadin, C. Fehn et al., “Depth map creation and image-based rendering for advanced 3d tv services providing interoperability and scalability,” *Signal Processing: Image Communication*, vol. 22, no. 2, pp. 217–234, 2007.

- [5] D. Sebai, "Performance analysis of hevc scalable extension for depth maps," *Journal of Signal Processing Systems*, vol. 92, no. 7, pp. 747–761, 2020.
- [6] C. Fehn, "Depth-image-based rendering (dibr), compression, and transmission for a new approach on 3d-tv," *SPIE*, vol. 5291, pp. 93–104, 2004.
- [7] E. G. Mora, J. Jung, M. Cagnazzo, and B. Pesquet-Popescu, "Initialization, limitation, and predictive coding of the depth and texture quadtree in 3d-hevc," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 9, pp. 1554–1565, 2014.
- [8] H. Hamout and A. Elyousfi, "An efficient edge detection algorithm for fast intra-coding for 3d video extension of hevc," *Journal of Real-Time Image Processing*, vol. 16, no. 6, pp. 2093–2105, 2017.
- [9] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (hevc) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.
- [10] H. Zhang, Y. Chan, C. Fu, S. Tsang, and W. Siu, "Quadtree decision for depth intra coding in 3d-hevc by good feature," in *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1481–1485, Shanghai, China, March 2016.
- [11] K. Peng, J. Chiang, and W. Lie, "Low complexity depth intra coding combining fast intra mode and fast cu size decision in 3d-hevc," in *Proceedings of the 2016 Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pp. 1126–1130, Phoenix, AZ, USA, September 2021.
- [12] 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)," *Journal of Visual Communication and Image Representation*, vol. 48, pp. 329–339, 2017.
- [13] M. Kim, N. Ling, and L. Song, "Fast single depth intra mode decision for depth map coding in 3d-hevc," in *Proceedings of the 2015 Proceedings of the IEEE International Conference on Multimedia Expo Workshops*, pp. 1–6, ICMEW, Turin, Italy, July 2019.
- [14] J. C. Chiang, K. K. Peng, C. C. Wu, C. Y. Deng, and W. N. Lie, "Fast intra mode decision and fast cu size decision for depth video coding in 3d-hevc," *Signal Processing: Image Communication*, vol. 71, pp. 13–23, 2019.
- [15] G. He, J. Hu, Y. Li et al., "Fast mode decision and pu size decision algorithm for intra depth coding in 3d-hevc," *Journal of Visual Communication and Image Representation*, vol. 49, pp. 303–314, 2017.
- [16] H. Hamout and A. Elyousfi, "Fast 3d-hevc pu size decision algorithm for depth map intra-video coding," *Journal of Real-Time Image Processing*, vol. 17, no. 5, pp. 1285–1299, 2019.
- [17] L. Shen, K. Li, G. Feng, P. An, and Z. Liu, "Efficient intra mode selection for depth-map coding utilizing spatiotemporal, inter-component and inter-view correlations in 3d-hevc," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4195–4206, 2018.
- [18] H. Yang, L. Shen, X. Dong, Q. Ding, P. An, and G. Jiang, "Low-complexity ctu partition structure decision and fast intra mode decision for versatile video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 6, pp. 1668–1682, 2020.
- [19] X. Dong, L. Shen, M. Yu, and H. Yang, "Fast intra mode decision algorithm for versatile video coding," *IEEE Transactions on Multimedia*, vol. 24, pp. 400–414, 2022.
- [20] M. S. Yang and C. Y. Lai, "A robust automatic merging possibilistic clustering method," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 1, pp. 26–41, 2011.
- [21] A. J. Díaz-Honrubia, J. De Praeter, J. L. Martínez, P. Cuenca, and G. Van Wallendael, "Reducing the complexity of a multiview h.264/avc and hevc hybrid architecture," *Journal of Signal Processing Systems*, vol. 90, no. 2, pp. 249–258, 2018.
- [22] V. Sze, M. Budagavi, and G. J. Sullivan, *High Efficiency Video Coding (Hvc)*, Springer International Publishing, Berlin, Germany, 2014.
- [23] A. Öztekin and E. Erçelebi, "An early split and skip algorithm for fast intra cu selection in hevc," *Journal of Real-Time Image Processing*, vol. 12, no. 2, pp. 273–283, 2016.
- [24] J. Lainema, F. Bossen, W. J. Han, J. Min, and K. Ugur, "Intra coding of the hevc standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1792–1801, 2012.
- [25] L. Shen, Z. Zhang, and Z. Liu, "Effective cu size decision for hevc intracoding," *IEEE Transactions on Image Processing*, vol. 23, no. 10, pp. 4232–4241, 2014.
- [26] H. Azgin, E. Kalali, and I. Hamzaoglu, "A computation and energy reduction technique for hevc intra prediction," *IEEE Transactions on Consumer Electronics*, vol. 63, no. 1, pp. 36–43, 2017.
- [27] Y. Piao, J. Min, and J. Chen, *Encoder Improvement of Unified Intra Prediction*, JCT-VC Document JCTVC-C207, Morocco, Africa, 2010.
- [28] E. Ozcan, E. Kalali, Y. Adibelli, and I. Hamzaoglu, "A computation and energy reduction technique for hevc intra mode decision," *IEEE Transactions on Consumer Electronics*, vol. 60, no. 4, pp. 745–753, 2014.
- [29] H. Hamout and A. Elyousfi, "Fast depth map intra coding for 3d video compression based tensor feature extraction and data analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 7, pp. 1933–1945, 2020.
- [30] J. C. Dunn, "A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters," *Journal of Cybernetics*, vol. 3, no. 3, pp. 32–57, 1973.
- [31] R. Krishnapuram and J. M. Keller, "A possibilistic approach to clustering," *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 2, pp. 98–110, 1993.
- [32] H. Hamout and A. Elyousfi, "Fast depth map intra-mode selection for 3d-hevc intra-coding," *Signal, Image and Video Processing*, vol. 14, no. 7, pp. 1301–1308, 2020.
- [33] D. H. R. Software, "Htm version 16.2," 2016, <https://hevc.hhi.fraunhofer.de/trac/3d-hevc/browser/3DVCSoftware/tags/HTM-16.2>.
- [34] K. Müller and A. Vetro, "Common test conditions of 3dv core experiments," in *Proceedings of the 7th Meeting Joint Collaborative Team on 3D Video Coding Extension Development Document JCT3V-G1100*, San Jos, CA, USA, June 2014.
- [35] G. Bjontegaard, "Calculation of average psnr differences between rd-curves," in *Proceedings of the 13th ITU-T SG 16 Q.6 VCEG, Document VCEG-M33*, Austin, Texas, USA, April 2001.