

# Fast intra coding algorithm for HEVC based on depth prediction and mode reduction

Fen Chen<sup>1</sup> · Defu Jin<sup>1</sup> · Zongju Peng<sup>1</sup> · Gangyi Jiang<sup>1</sup> · Mei Yu<sup>1</sup> · Hua Chen<sup>1</sup>

Received: 25 June 2017 / Revised: 1 March 2018 / Accepted: 16 April 2018 /  
Published online: 30 April 2018  
© Springer Science+Business Media, LLC, part of Springer Nature 2018

**Abstract** The latest High Efficiency Video Coding (HEVC) standard only requires 50% bitrate of the H.264/AVC at the same video quality. Due to introduction of more advanced coding tools and techniques, its computational complexity rises rapidly. In this paper, we propose a novel fast Intra coding algorithm. Firstly, a novel feature is proposed to measure the complexity of video content based on the analysis of mode information obtained from a previous frame. Then, a model is built based on the relationship between this feature and Coding Tree Unit (CTU) depth range. According to the model, the unnecessary operations of Coding Unit (CU) split are skipped. Secondly, the correlation between CU distribution structure and mode selection is established. For the prediction of unlikely CU, the last few intra modes in candidate list are eliminated before rate-distortion optimization process. Experimental results demonstrate that the proposed algorithm reduces the encoding time by 43.2% with just 0.47% BD-rate increase compared with HEVC test model. Compared with other state-of-art algorithms, the proposed algorithm achieves better trade-off between complexity reduction and RD performance.

**Keywords** HEVC · Intra coding · Prediction mode · CU distribution structure

## 1 Introduction

With the development of capture and display technologies, High Definition (HD) video is gradually applied into many fields, such as, television, games, education, virtual reality and augmented reality [37, 38]. The data amount of HD video is tremendous, which exceeds the network transmission bandwidth. It is critical to develop efficient compression technologies. The most well-known compression methods are the Joint Photographic Experts Group (JPEG)

---

✉ Zongju Peng  
pengzongju@126.com

<sup>1</sup> Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang, China

and JPEG 2000 [7–9]. These methods are based on transform coding and perform better than others. However, they are only suitable for image. To efficiently store and transmit video signal, many video compression standards such as MPEG4, H.264/AVC [28] are successively proposed. Recently, ISO and ITU have jointly established a new video coding standard, High Efficiency Video Coding (HEVC) [35]. Compared with traditional standard H.264/AVC, HEVC achieves half bitrate saving for high definition video with same reconstructed video quality.

In HEVC, intra coding [12] is employed to exploit the spatial redundancy. Compared with H.264/AVC standard, HEVC intra coding not only introduces the quad-tree-based recursive partitioning and fine-grain intra prediction technology, but also adopts advanced Coding Unit (CU), Prediction Unit (PU) and Transform Unit (TU) [40]. Although the use of these new technologies and coding tools provides high flexibility for encoding the video with smooth or complex content, it also brings huge encoding complexity, which hinders the real-time application of HEVC.

Many algorithms have been proposed to reduce the computational complexity for H.264/AVC, which can be roughly classified into two categories. The first category [11, 16, 36] reduces the impossible candidate modes by making full use of statistical Rate-Distortion (RD) cost, dominant edge direction and the spatial and temporal correlation. The second category [10, 14, 15] reduces the computational complexity of motion estimation. However, these methods cannot be directly applied to HEVC due to the introduction of new coding tools and technologies.

Recently, a number of fast algorithms have been proposed to reduce the computational complexity for H.265/HEVC. These algorithms can be classified into two types, CU depth early termination or skip and fast intra mode decision. For CU depth early termination or skip, numerous algorithms have been developed by analyzing the texture complexity of video content, spatiotemporal depth correlation and RD cost. In [34], the current CU is firstly classified as homogeneous or complex according to texture homogeneity measured by mean absolute deviation. Then, the classification result and coding information from neighboring coded CUs are used to determine the optimal CU size. In [30], the CU smoothness parameter is generated based on the variance of the SADs in horizontal, vertical, left and right diagonal directions. According to the variance, the CU split process is early terminated or skipped based on two predefined thresholds. In [26], the global and local edge complexity in horizontal, vertical, left and right diagonal directions are proposed and used for CU splitting. If the global and local edge complexity are lower than predefined thresholds decided by the wrong hit rate, the searching of CU split is early terminated. In [2], early termination of simple CTU is achieved according to the edge information detected by Sobel operator. In [24], the unnecessary coding block will be skipped adaptively by calculating the texture complexity for each coding depth. In [25], an adaptive double thresholds scheme is employed to classify the video content. Then, the classification and depth information of temporally co-located CU and the spatially neighboring CUs are jointly utilized to constrain the CTU depth range. In [33], some specific depth levels rarely used in spatially neighboring CUs are skipped. In [32], the depth information of spatially neighboring CUs is exploited to make an early CU splitting decision or CU pruning decision. In [43], based on the temporal correlation, the depth information of temporally co-located CTU is used to predict the size of current CTU. In [39], the aggregated RD cost of sub-CUs is estimated according to Hadamard cost and RD cost of partial sub-CUs. If the aggregated RD cost is larger than the RD cost of current CU, the CU split process is early terminated. In [5], a rough mode cost (RMC) based algorithm is proposed for early CU

splitting or termination. Whether to split or terminate the current CU is decided by comparing the least RMC value with the adaptive threshold.

For fast intra mode decision, related works have been proposed by analyzing the CU texture direction, RD cost and spatial mode correlation. In [24], the pixel values deviation technique is utilized to obtain the dominate texture direction for each PU. Based on the dominate direction, the number of candidate modes for Rough Mode Decision (RMD) and RD Optimization (RDO) is reduced from 35 to 13 and 8 to 3, respectively. In [33], the RD cost and prediction mode correlations among different depth levels or spatially nearby CUs are exploited to reduce some prediction modes. In [32], the prediction mode correlation between the higher layer PU and current layer PU is utilized to terminate some prediction modes which are rarely selected as the optimal mode. In [43], the dominant direction of the block is detected by discrete cross difference to reduce the prediction modes for RMD process. In addition, four early termination methods are proposed to terminate the RDO process properly. In [39], a progressive rough mode search method is proposed to check the potential modes instead of traversing all modes. Besides, an early RDO skip scheme is introduced for further complexity reduction. If the mode distance between current mode and any evaluated mode is lower than 2, the current mode is excluded for RDO process. In [5], the ratios between the least RMC value and all other RMC values are employed to reduce the number of intra modes for RDO process. Based on the correlation between Hadamard cost and RD cost, a selective function is established to reduce some modes for RDO process in [17]. In [31], the dominant texture orientation is detected by computing the local directional variance along a set of co-lines. According to the dominant texture orientation, a reduced number of candidate modes are selected to be further checked by RDO process. In [27], the prediction modes for RMD are decreased based on mode correlation among upper CU or spatially nearby PUs. For further complexity reduction, an early termination method for RD optimization (RDO) and RQT processes is introduced by an adaptive threshold based on the average value of RD costs. In [18], a two-step fast mode decision algorithm is proposed for HEVC. Firstly, the depth information is used to skip some intra modes for RMD process. Then, the order of RMD and MPM is adjusted to skip some unnecessary check for RDO. In [1], the texture of block diagonals is used to reduce the number of modes for RMD process.

In addition to traditional statistical methods mentioned above, many learning methods are extensively used in prediction and recognition [19–23, 29]. Some advanced learning algorithms are introduced for fast video coding. In [41, 42, 44], the authors modelled CU partitioning and mode selection as a binary or multi-class problem. Then, joint Support Vector Machine (SVM) based classifier is applied to early CU termination, early CU split decision and fast mode selection. In [13], Kim et al. propose a fast CU size decision algorithm, which early terminates the CU split process based on Bayesian decision rule. In [6], a decision tree is built through data mining techniques to decide whether the CU split process should be early terminated or not. For these learning algorithms, multiple features can be jointly used. Besides, sophisticated learning algorithms can improve the prediction accuracy when compared with traditional statistical methods.

For the content complexity-based fast CU size algorithm proposed in previous work, the variance or Sobel gradient are used to constrain the depth range of CTU. These algorithms fail on video content with single direction texture, which can be predicted well by corresponding angular mode. For the correlation based fast CU size algorithm mentioned above, the depth information of spatially neighboring CUs or co-located CU is exploited to accelerate the CU split process. However, these algorithms do not make full use of the intra prediction

mechanism characteristic in HEVC. After RMD process, some methods are proposed to reduce the number of modes checked by RDO process based on the information of mode candidate list. These early termination algorithms could make the optimal prediction mode skipped, which leads to considerable loss of RD performance.

To alleviate abovementioned drawbacks, we propose a simple and effective intra coding algorithm based on depth range prediction and mode reduction. In HEVC intra coding, up to 35 modes are used for prediction, and each of them can be used to predict a block with single direction texture. There exists implicit relationship between intra modes and video texture characteristics. After encoding, it can be inferred that the texture of current block is rich if various intra modes are included. Based on this observation, instead of variance and Sobel gradient, a novel feature is proposed to measure the complexity of video content. Then, the relationship between this feature and CTU depth range is modeled. According to the model, the unnecessary CU split operations are skipped. After RMD process, the candidate modes in list are sorted in descending order according to Hadmard cost. That is, the modes rank ahead have high probability to be selected as optimal mode. Conversely, the modes rank behind in candidate list are rarely chosen as best mode. For each CU, the prediction is conducted only when the split flag of parent CU is set true. Based on aforementioned analyses, a fast mode decision algorithm is proposed to reduce the encoding complexity. For the prediction of unlikely CU, the last few intra modes in candidate list are eliminated before RDO process. Experiment result shows that the proposed overall algorithm can save significant encoding time reduction with negligible loss of RD performance, achieving well trade-off between complexity reduction and RD performance. This paper has two main contributions. 1) A novel feature is proposed to measure the complexity of video content. Based on the analysis about the feature, we build a depth prediction model for CTU. 2) We propose a fast mode decision algorithm for RDO process based on the statistical analyses of CU distribution structure.

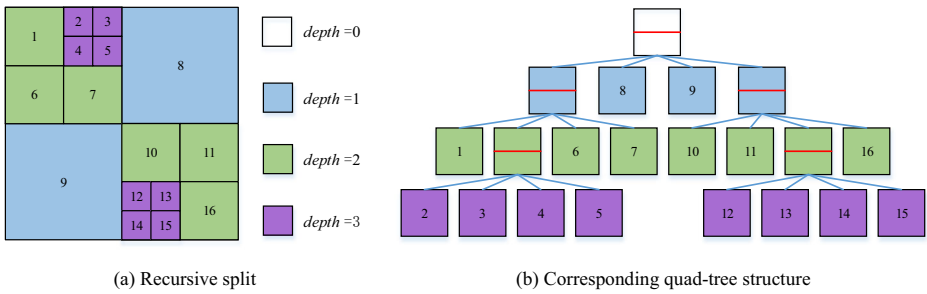
The rest of this paper is organized as follows. In section 2, the procedure and its complexity analysis of HEVC intra coding is presented. Section 3 presents a fast intra coding algorithm based on correlation between PU prediction and CU split. The experimental result and conclusion are illustrated in Section 4 and Section 5, respectively.

## 2 HEVC intra coding and complexity analysis

To improve compression performance, HEVC introduces many novel technologies with the cost of rapid increase in computational complexity at encoder side. The significant computational load for intra coding is primarily due to its quad-tree-based recursive partitioning and fine-grain intra prediction [5]. The retailed analysis is presented in this section.

Compared to traditional encoding standard, HEVC still adopts block-based encoding style. Firstly, each frame is divided into non-overlapping CTUs. Then, flexible quad-tree structure is employed to recursively split each CTU into different size CUs. The size of CU can be  $64 \times 64$ ,  $32 \times 32$ ,  $16 \times 16$  and  $8 \times 8$ , corresponding to depth of 0, 1, 2 and 3. In CU split process, each CU can be further split into one or multiple PUs and TUs by prediction and transform.

In HEVC, the CU split and mode decision are determined by the RDO rule, which represents a trade-off between the bit-rate and reconstructed quality. Aiming at achieving the best RD performance, HM encoder traverses every possible CU combination to get the optimal split scheme which leads to minimum RD cost. Figure 1 shows the recursive split of a CTU. Block with different color represents different depth CUs. The number and red line inside each



**Fig. 1** Diagram for CTU split

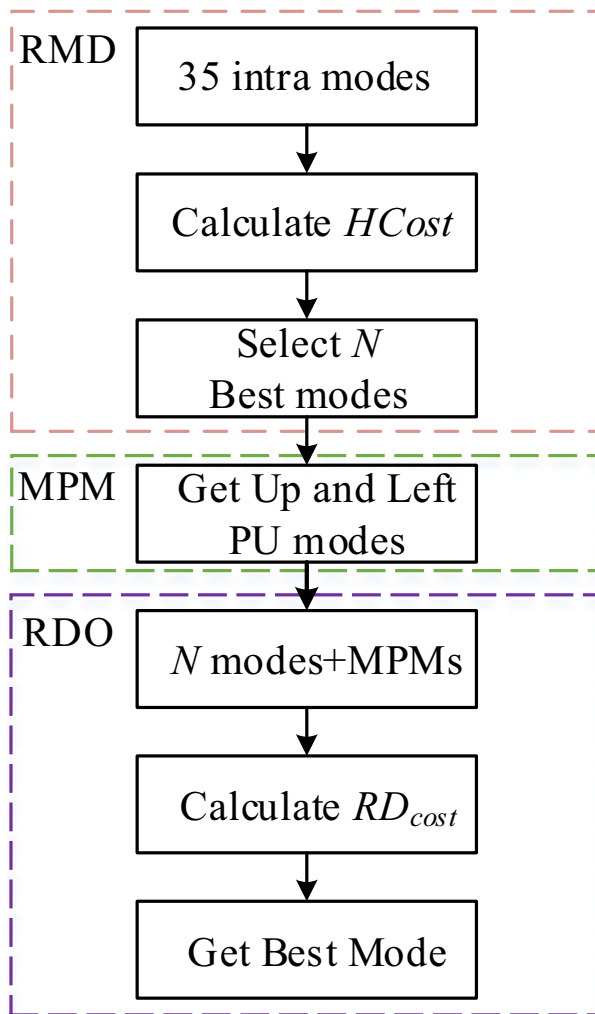
CU represents the traversing order and split flag, respectively. The specific split process can be described in two steps. Firstly, traverse every depth CU by the Z-scan order and calculate RD cost. And then, backward pruning is performed according to the RD cost until the coding structure of CTU is determined.

Intra prediction is executed to get the optimal intra mode by encoder at each depth of CTU. It includes RMD, Most Probable Mode (MPM) derivation and RDO process, as show in Fig. 2. Firstly, RMD is performed to select  $N$  candidate modes according to the Hadamard cost, where the value of  $N$  is set 3, 3, 3, 8 and 8 for  $64 \times 64$ ,  $32 \times 32$ ,  $16 \times 16$ ,  $8 \times 8$  and  $4 \times 4$  PU, respectively. After this process, the number of intra modes for RDO is reduced from 35 to  $N$ . Then, additional MPMs from left and above PUs will be added into candidate list if they are not included yet. Due to high correlation of spatially nearby PUs, this process can improve the coding efficiency. In the end, full RDO process is executed to choose the optimal mode from candidate list.

With the above analyses, the decision of coding parameters for each CTU requires encoder traversing different depth CUs, which needs  $85(1 + 4 + 16 + 64)$  recursions. At each depth of CTU, PU prediction is executed to get the optimal intra mode. At depth 3, PU can be further divided into four sub-PUs. HEVC intra coding supports 35 prediction modes, including DC, Planar and 33 angular modes. For luminance component, 11,935 rough mode predictions for RMD process and at least 2623 refined mode predictions for RDO process need to be performed to determine the optimal mode for each CTU [24]. And the RDO process includes a series of complex processes such as prediction, transform, quantization and reconstruction. Hence, the introduction of new coding tools and technologies in HEVC requires massive computing resources, which blocks the practical encoding applications.

### 3 Proposed algorithm

In order to reduce the computational complexity for HEVC intra coding, an optimization algorithm is proposed. Figure 3 shows the overall flowchart of proposed algorithm, where  $C$  is current CTU and  $DL$  is the CU depth level. The overall algorithm includes Mode-based CTU Depth Range Prediction (MCDRP) and Distribution-based Intra Mode Fast Selection (DIMFS). The MCDRP is used to predict probable depth range of current CTU based on mode information obtained from a previous frame. For each depth, the DIMFS algorithm is performed to reduce the candidate modes for RDO process. Before encoding a CTU, the proposed MCDRP algorithm is performed to determine the  $DL$  range. The  $DL$  range of original HM14.0 platform is from 0 to 3. After depth prediction, the  $DL$  range is from min to max. For each depth of  $DL$  range, the DIMFS algorithm is recursively conducted. Since the



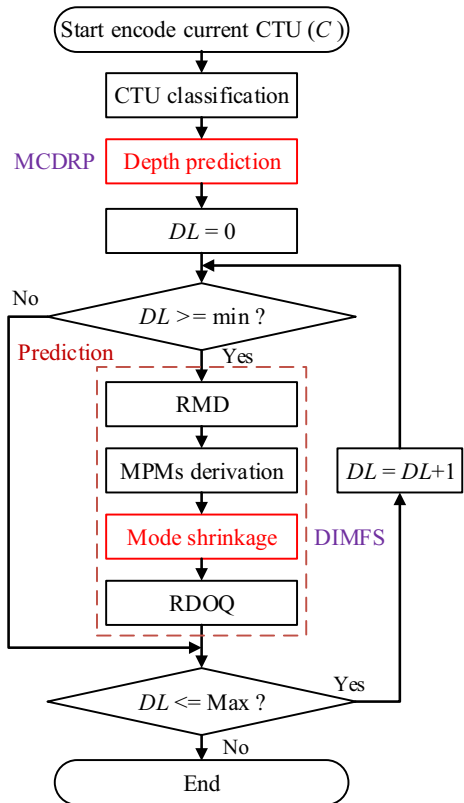
**Fig. 2** Intra prediction process

two algorithms work for different stage, the overall algorithm can achieve considerable encoding time reduction.

### 3.1 Mode-based CTU depth range prediction algorithm

In HEVC, CTU split process needs to traverse all depth CUs, which exists large depth redundancy for CTUs whose maximum split depth is only 0, 1 and 2. The unnecessary split operations will be reduced if the depth range of CTU can be predicted before encoding. This paper utilizes the mode information of co-located CTU from a previous frame to predict the depth range of current CTU.

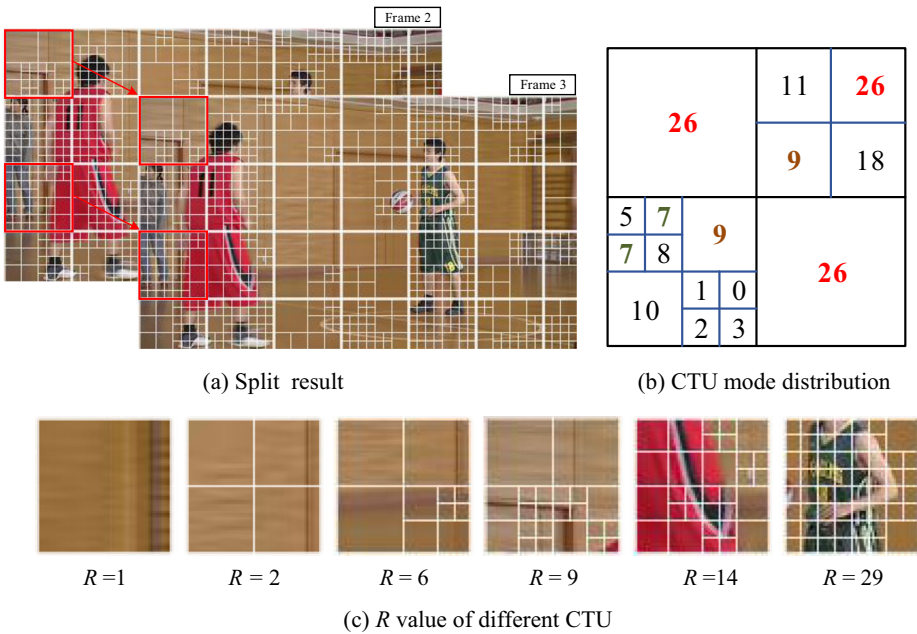
**Fig. 3** Overall flowchart of proposed algorithm



Because the contents between successive frames of each video are almost the same, the encoding parameters adopted by temporally co-located CTU and current CTU are similar [43]. For example, Fig. 4(a) shows the optimal split result for the second and third frames of BasketballPass sequence when QP is 32. As shown in the figure, the split depth between co-located CTUs is quite similar. HEVC adopts 35 prediction modes for intra coding, and each of them represents a directional texture. It is obvious that one CTU includes rich texture if many prediction modes are used. For nature video, the split structure of the region with rich texture tends to be complex. The number of different modes used for a CTU is defined as  $R$ . Figure 4(b) shows the mode distribution for a CTU with  $R$  equals to 12, where the number inside CUs represents a prediction mode. Figure 4(c) lists the six CTUs and its  $R$  value from the third frame of BasketballPass sequence. With the increase of  $R$  value, the split structure of CTU is more complex. Above analysis suggests that the  $R$  value of co-located CTU from a previous frame can be used to predict the depth range of current CTU.

It is obvious that the smooth region is more inclined to be coded by large CUs. In contrast, the region with rich texture is more likely to be coded into small CUs. Firstly, this paper classifies different CTUs into three categories according to depth information. A CTU is classified as simple CTU (SCTU) when all of CUs take a depth no larger than 1 and as complex CTU (CCTU) when all of CUs take a depth no smaller than 2. All other CTUs are defined as median CTU





**Fig. 4** Encoding parameters for the second and third frames for BasketballPass sequence

(MCTU). It is noted that the SCTU has the maximum  $R$  value 4 when all of CUs inside CTU take depth 1 and the CCTU has the maximum  $R$  value 16 when all of CUs inside CTU take depth 2. Then, the relation between  $R$  and CTU Complexity Type ( $CT$ ) can be modeled as

$$CT = \begin{cases} \text{SCTU}, & 1 \leq R \leq 4 \\ \text{MCTU}, & 4 < R < 16 \\ \text{CCTU}, & 16 \leq R \leq 35 \end{cases} \quad (1)$$

To analyze the relation between the practical split depth and classification results, the optimal depth information of SCTU, MCTU and CCTU is investigated respectively, as shown in Table 1. The  $nd3$  column represents the percentage of CU whose depth smaller than 3 while the  $nd0$  and  $nd01$  columns show the percentage of CU with depth not less than 1 and depth not less than 2. The average value of  $nd3$  is 93.1% for SCTU. For MCTU and CCTU, the average value of  $nd0$  is 98.0% and 100.0%, respectively. Furthermore, the average value of  $nd01$  for CCTU is 89.9%. That is, SCTU is more inclined to be coded by large CUs. Conversely, the small CUs are better for MCTU and CCTU. A state-of-art depth prediction algorithm is proposed by Zhao [43], which the depth information of co-located CTU is used. The algorithm is implemented for comparison under the same condition, as shown in Table 1. Whether the CTU belongs to SCTU or CCTU, the accuracy of proposed algorithm is higher than Zhao’s algorithm. To reduce encoding complexity as much as possible, the relation between the  $CT$  and depth range ( $DR$ ) is modeled as

$$DR = \begin{cases} [0, 2], & CT \in \text{SCTU} \\ [1, 3], & CT \in \text{MCTU} \\ [2, 3], & CT \in \text{CCTU} \end{cases} \quad (2)$$



**Table 1** Optimal depth distribution of different CTUs (%)

Sequence	Proposed				Zhao [43]	
	SCTU		CCTU		SCTU	CCTU
	<i>nd3</i>	<i>nd0</i>	<i>nd0</i>	<i>nd01</i>	<i>nd3</i>	<i>nd0</i>
BQMall	83.3	99.4	100.0	95.1	79.1	99.8
Johnny	98.5	99.1	100.0	90.8	98.0	99.4
Kimono	98.8	94.6	100.0	79.6	96.8	96.5
Traffic	92.0	98.8	100.0	94.2	82.9	99.2
Average	93.1	98.0	100.0	89.9	89.2	98.7

In this way, the average accuracy of proposed algorithm can reach to 89.9%, guaranteeing the quality of video.

### 3.2 Distribution-based intra mode fast selection

Refined mode prediction process is one of main responsible for the increase of encoder's complexity. Although the predefined 35 modes are reduced to 3 or 8 by RMD algorithm, the encoding complexity is still high and there exists some redundancy. In HEVC intra coding, a PU partitioning structure has its root at the respective CU level. This means that the PU prediction depends on the CU structure. If the split structure of CTU is known, the unnecessary predictions can be skipped without changing the optimal result. For example, if the *DR* value of one CTU is [9, 38], encoder only needs to traverse the depth from 2 to 3. In this section, we proposed to reduce the complexity of prediction process by exploring the correlation between CU structure and mode prediction.

Using the HEVC test model HM14.0, 5 different resolution sequences from class A to class E given by common test conditions [4] are tested to analyze the CU distribution structure, as shown in Table 2. The average percentage of CU from 0 to 3 depth is 8.12%, 25.64%, 26.67% and 39.58%, respectively. That is, the 0 and 1 depth CUs have a high probability to be split into smaller CUs. For CU at 3 depth, there is only 39.58% percentage to be selected as the optimal depth. There is no need to check all candidate modes for RDO process when the current CU has a lower probability to be selected as

**Table 2** Statistics of optimal CU structure (%)

CU Size	64 × 64	32 × 32	16 × 16	8 × 8
class A	4.54	22.47	35.73	37.26
class B	13.57	34.48	24.61	27.35
class C	1.00	18.02	25.69	55.29
class D	0.74	17.83	27.40	54.03
class E	17.16	32.22	28.98	21.65
Average	8.12	25.64	26.67	39.58

optimal depth. After RMD process, the modes inside candidate list are sorted by Hadamard cost in descending order. The distribution of Hadamard cost of a PU is highly correlated with the distribution of RD cost [17]. This implies that the mode ranked behind in candidate list has a low probability to be selected as the best mode by RDO process.

Based on above analysis, a simple and effective algorithm is proposed to reduce the number of candidate mode for RDO process. Firstly, the prediction accuracy of different size PU is calculated when adjusting the  $N$  value, as shown in Fig. 5. And then, we propose revising the accuracy by taking the CU structure into consideration. We define a set  $\psi = \{mode_1, mode_2, \dots, mode_N\}$  represents the first  $N$  modes in the candidate list. The original accuracy and revised accuracy are denoted by  $P$  and  $P_r$ , and they are calculated as shown in Eqs. (3) and Eqs. (4), respectively. Where  $n_N$  is the number of PU whose best prediction mode belongs to  $\psi$  and  $n_{HM}$  is the total number of PU. The *factor* is a probability related to CU distribution. When the PU size is  $64 \times 64$ , the CU has a probability of 91.88% to be split. That means, the wrong predictions have the same probability not to influence the final prediction result. Therefore, the factor value of  $64 \times 64$  size PU is set as 0.9188. Similarly, the *factor* value of  $32 \times 32$ ,  $8 \times 8$  and  $4 \times 4$  size PU are 0.6625, 0.6142 and 0.6142, respectively.

$$P = \frac{n_N}{n_{HM}} \tag{3}$$

$$P_r = P + (1-P) \times factor \tag{4}$$

As shown in Fig. 5, the accuracy can be improved a lot after revising. In the mode prediction process, the number of candidate mode can be reduced by eliminating such structural redundancy. Under the condition that the revised accuracy is less than 3%, the  $N$  value of  $64 \times 64$ ,  $32 \times 32$ ,  $8 \times 8$  and  $4 \times 4$  size PU are sequentially set to 1, 2, 4, and 4. After processing, the calculations for the RD cost of each CTU is reduced from 2623 to  $1337(1 + 2 \times 4 + 3 \times 16 + (64 + 256) \times 4 = 1337)$ , which significantly reduces the complexity of RDO

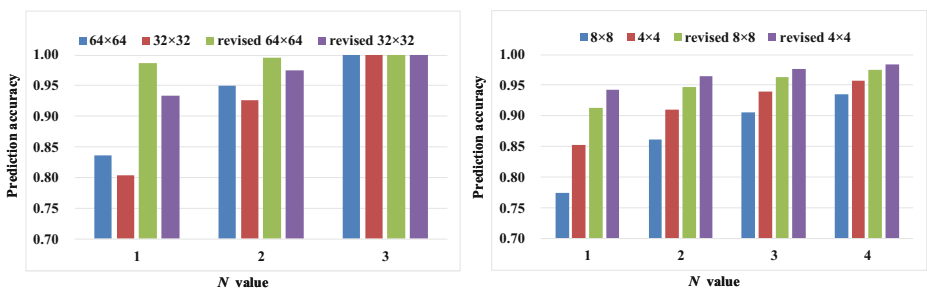


Fig. 5 Mode prediction accuracy when adjusting  $N$  value

process. Compared with  $32 \times 32$  PU, the *factor* for  $16 \times 16$  PU is lower than 40%, and for this reason the proposed algorithm is not applied for  $16 \times 16$  PUs.

---

**Algorithm 1** Fast CU size and mode decision for HEVC intra coding

---

```

1: define  $S_i$  represent the number of modes for RDO process
2:  $S = \{1, 2, 3, 4\}$ 
3: if  $R \leq 4$  then
4:    $\min\_depth = 0, \max\_depth = 2$ ;
5: else if  $R > 4 \ \&\& \ R \leq 16$ 
6:    $\min\_depth = 1, \max\_depth = 3$ ;
7: else
8:    $\min\_depth = 2, \max\_depth = 3$ ;
9: end if
10: for  $\min\_depth \leq i \leq \max\_depth$  do
11:   for  $0 \leq j \leq 35$  do
12:     // perform RMD process
13:   end for
14:   //perform MPM
15:   for  $0 \leq k \leq S_i$  do
16:     // perform RDO process
17:   end for
18: end for

```

---

According to the above analyses, the proposed whole algorithm could be summarized in Algorithm 1. Firstly, the  $R$  value of current CTU is calculated based on the mode information obtained from a previous frame. Then, the  $\min\_depth$  and  $\max\_depth$  of current CTU are determined according to Eqs. (1) and (2). On the basis of proposed DIMFS algorithm, the number of modes for RDO process is predetermined. For CU depth from 0 to 3, the number of modes for RDO process is set 1, 2, 3 and 4, respectively. In the end, the prediction is recursively conducted at CU depth from  $\min\_depth$  to  $\max\_depth$ .

## 4 Experimental results

To verify the performance of the proposed CU size and PU mode decision algorithms, we implemented them on the HEVC reference platform HM 14.0 with the All-intra-main configuration. All the coding experiments are performed on the computer with Inter(R) Core(TM) i5–4590 CPU @3.3GHz and 8GB RAM, Windows 7 64-bit operating system. Four QPs, 22, 27, 32 and 37 are used. The Size of largest CU and smallest CU are  $64 \times 64$  and  $8 \times 8$ , respectively. The GOP size is 1 and the other coding parameter settings follow the definitions in common test condition [4]. In addition, encoding time is used to measure the computational

complexity, and time saving ratio ( $\Delta T$ ) is employed to measure the complexity reduction. RD performance is measured with BD-rate, which is defined in [3].  $\Delta T$  is defined as.

$$\Delta T(\%) = \frac{1}{4} \sum_{i=1}^4 \frac{T_{HM}^{OP_i} - T_{Proposed}^{OP_i}}{T_{HM}^{OP_i}} \times 100, \underline{OP}_i = \{22, 27, 32, 37\} \tag{5}$$

where  $T_{HM}^{OP_i}$  and  $T_{Proposed}^{OP_i}$  is the encoding time of original HM and proposed algorithm with  $OP_i$ , respectively.

Twenty standard video sequences with various resolutions and contents, as shown in left part of Table 3, are encoded. Table 3 shows the encoding time reduction and RD performance of proposed algorithm. Compared with HM14.0, the proposed DIMFS and MCDRP algorithms can respectively save encoding time by 18.5% and 31.5% on average with just 0.23% and 0.24% BD-rate increase. The encoding time reduction of different sequence for MCDRP algorithm exists obvious difference and for DIMFS algorithm is almost the same. There are two mainly reasons for this phenomenon. On the one hand, the strategy adopted by the DIMFS algorithm between different sequences is same due to the global CU structure feature. On the other hand, the feature adopted by MCDRP algorithm is related to the video content. The video content of the partial sequence like Kimono is relatively simple and the CTU type is mainly SCTU. For such CTU, the complex prediction process of  $8 \times 8$  and  $4 \times 4$  PU will be early terminated according to 3.1 section. Therefore, the encoding time reduction for this kind of sequence is significant. The proposed overall algorithm can save encoding time by 43.2% on average with just a 0.47% BD-rate increase.

To intuitively compare the RD performance of proposed algorithm and original HM14.0 algorithm, we draw the RD curves and local enlargements of “BasketballDrive” and “Kimono” which are shown in Fig. 6. The blue and red curves respectively represent the RD performance

**Table 3** RD performance and encoding time reduction of the proposed algorithm compared to HM14.0

Class	Sequence	Frame rate(fps)	Bit depth	DIMFS (%)		MCDRP (%)		Proposed (%)	
				BD-rate	$\Delta T$	BD-rate	$\Delta T$	BD-rate	$\Delta T$
<b>A</b> 2560 × 1600	Traffic	30	8	0.24	18.1	0.19	23.7	0.43	37.2
	PeopleOnStreet	30	8	0.28	18.5	0.16	21.8	0.43	36.5
	Nebuta	60	10	0.07	18.7	0.17	38.5	0.23	48.0
	SteamLocomotive	60	10	0.05	18.7	0.11	46.8	0.17	54.0
<b>B</b> 1920 × 1080	Kimono	24	8	0.16	17.4	0.15	48.1	0.29	54.7
	ParkScene	24	8	0.21	18.4	0.23	26.0	0.46	39.2
	Cactus	50	8	0.27	18.5	0.20	25.8	0.45	39.3
	BasketballDrive	50	8	0.19	18.4	0.95	34.9	1.14	45.4
<b>C</b> 832 × 480	BQTerrace	60	8	0.21	19.4	0.12	27.6	0.36	41.2
	BasketballDrill	50	8	0.19	18.2	0.22	19.4	0.41	34.5
	BQMall	60	8	0.31	18.6	0.25	22.7	0.56	37.7
	PartyScene	50	8	0.42	18.5	0.02	22.8	0.46	38.4
<b>D</b> 416 × 240	RaceHorses	30	8	0.19	18.3	0.26	20.8	0.42	35.7
	BasketballPass	50	8	0.23	18.7	0.31	22.2	0.56	36.4
	Flowervase	50	8	0.35	18.5	0.13	29.8	0.49	41.8
	BlowingBubbles	50	8	0.33	18.7	0.01	19.8	0.35	35.9
<b>E</b> 1280 × 720	RaceHorses	30	8	0.29	18.2	0.14	19.9	0.43	35.2
	Vidyo1	60	8	0.25	17.9	0.25	28.6	0.52	40.5
	Vidyo3	60	8	0.22	18.5	0.59	31.3	0.77	42.9
	Vidyo4	60	8	0.19	18.2	0.33	27.6	0.53	40.1
<b>Average</b>				0.23	18.5	0.24	31.5	0.47	43.2

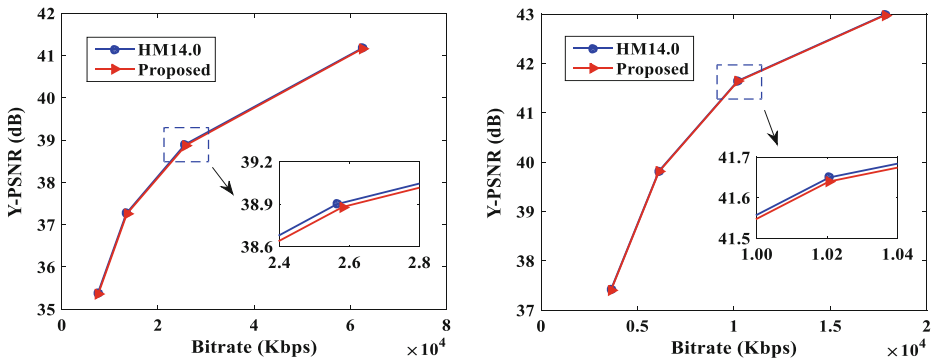


Fig. 6 RD curve of “BasketballDrive” and “Kimono”

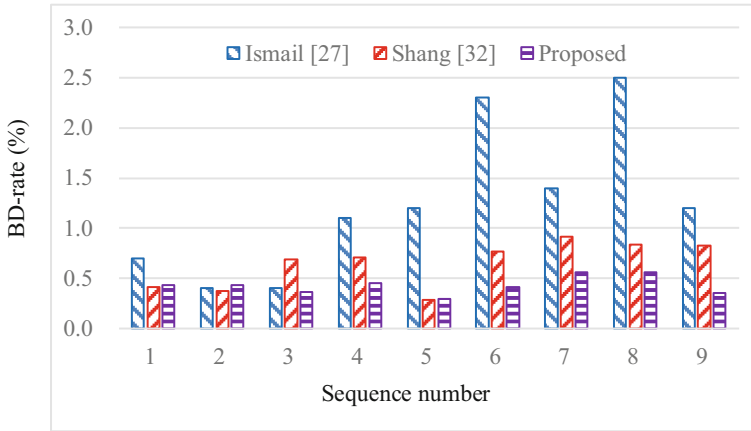
of original HM14.0 platform and our proposed algorithm. The two curves nearly overlap each other which implies that the RD performance of the proposed algorithm is almost the same with HM14.0 platform. In Table 3, the BD-rate increase of “BasketballDrive” is largest among all test sequences. Still, the RD performance degradation of ”BasketballDrive” is negligible. It proves that the proposed algorithm can guarantee the RD performance.

We compared the proposed algorithm with the state-of-art algorithms proposed by Shang [32] and Ismail [27]. As shown in Table 4, Ismail’s algorithm can save encoding time by 40.0% on average with 1.30% loss of BD-rate. Although additional threshold is employed to improve the accuracy, RDO early skip algorithm could make several optimal modes skipped, which results in significant loss of RD performance. Shang’s algorithm can save encoding time by 37.9% on average with only 0.66% loss of BD-rate. Compared with Ismail’s algorithm, the proposed algorithm can achieve 3% more time reduction. Meanwhile, the BD-rate loss of our algorithm is 0.83% lower than Ismail’s algorithm. Compared with Shang’s algorithm, the proposed algorithm can achieve more 5% time reduction. Meanwhile, the loss of BD-rate of our algorithm is 0.19% lower than Shang’s algorithm. Especially for large definition sequences including “Nebuta”, “SteamLocomotive”, “Kimono”, etc., the proposed algorithm can achieve 10% more time reduction than these reference algorithms.

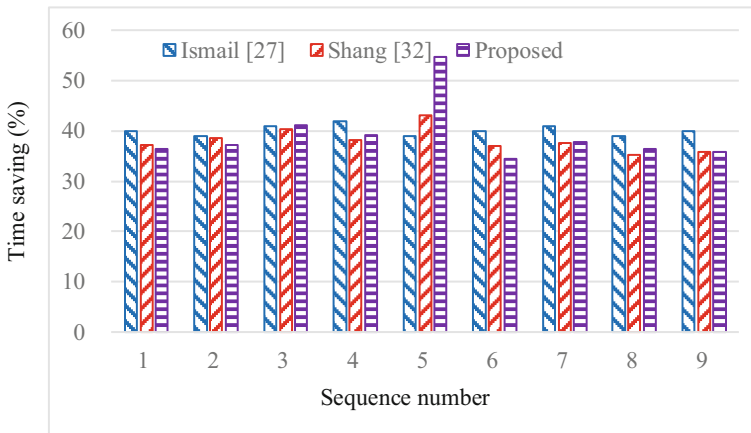
Figure 7 shows the BD-rate and time saving among Shang [32], Ismail [27] and the proposed algorithm. Nine sequences with different resolutions and contents are contained. The sequence number from 1 to 9 are “PeopleOnStreet”, “Traffic”, “BQTerrace”, “Cactus”, “Kimono”, “BasketballDrill”, “BQMall”, “BasketballPass”, “BlowingBubbles”, respectively.

Table 4 Performance comparison between proposed algorithm and other two state-of-art algorithms

Class	Shang [32]		Ismail [27]		Proposed	
	BD-rate (%)	$\Delta T$ (%)	BD-rate (%)	$\Delta T$ (%)	BD-rate (%)	$\Delta T$ (%)
A	0.39	37.8	0.57	39.2	0.31	44.1
B	0.56	40.5	0.76	40.6	0.54	43.5
C	0.76	37.1	1.22	40.0	0.46	36.7
D	0.87	35.0	2.08	38.7	0.46	37.0
E	0.65	40.0	1.97	40.7	0.61	41.2
Average	0.66	37.9	1.30	40.0	0.47	43.2



(a) BD-rate



(b) Time saving

**Fig. 7** BD-rate and time saving among reference algorithms and proposed algorithm

It is observed that the loss of BD-rate of proposed algorithm is lower than 0.60%. The RD penalty of the proposed algorithm is always minimal among three algorithms. For most sequences, the difference of time reduction among three algorithms is quiet small. For sequence Kimono, it is evident that the time reduction of proposed algorithm is better than other two algorithms. In addition, we can observe that RD penalty of Ismail’s algorithm is large for some sequences such as “BasketballDrill”, “BQMall”, “BasketballPass” and “BlowingBubbles”. This mainly because the threshold for RDO early skip algorithm is not always effective for all sequences. In the proposed algorithm, the strongly temporal correlation is exploited to constrain the depth range of each CTU, rather than to exhaustively search at each depth level. In the mode decision, some unlikely candidate modes are skipped, which make full use of the intra prediction mechanism characteristic. Therefore, compared to the benchmark, the encoding speed is accelerated, and the RD performance is guaranteed.

Besides the two classical algorithms above, we also select the two latest learning algorithms proposed by Zhang [41] and Zhang [42] for comparison, as shown in Table 5. It is observed

**Table 5** Performance comparison between the existing algorithms and the proposed algorithm

Algorithms	BD-rate (%)	$\Delta T$ (%)	BD-rate/ $\Delta T$
Shang [32] (HM14.0)	0.66	37.9	0.017
Ismail [27] (HM14.0)	1.30	40.0	0.032
Zhang [41] (HM14.0)	0.70	53.9	0.013
Zhang [42] (HM15.0)	1.58	52.8	0.030
Proposed (HM14.0)	0.47	43.2	0.011

that these two learning algorithms can respectively save 53.9% and 52.8% encoding time on average with 0.70% and 1.58% BD-rate penalty. In terms of time saving, the proposed algorithm is lower than learning algorithms, and superior to Ismail's algorithm and Shang's algorithm. For the comparison of RD performance, the BD-rate penalty of proposed algorithm is far superior to other algorithms. To make a fair comparison, the ratio of BD-rate to  $\Delta T$  proposed by Correa [6] is employed to measure the overall performance of various algorithms. It is obvious that the lower BD-rate/ $\Delta T$  ratio represents the better performance as it is interpreted as a good balance between low penalty and high speed up. It is observed that the BD-rate/ $\Delta T$  value of proposed algorithm is 0.011 which is least among these algorithms. Although the complexity reduction of proposed algorithm is still not as much as learning algorithms, the proposed algorithm can achieve the better trade-off between RD performance and complexity reduction than current state-of-art algorithms.

## 5 Conclusions

To reduce the complexity of HEVC, this paper proposes a simple and effective intra coding algorithm based on depth range prediction and mode reduction. In the CU depth early termination or skip, the number of different modes adopted by co-located CTU from a previous frame is employed to predict the depth range of current CTU. In the mode selection, CU distribution structure is employed to reduce the number of candidate mode for RDO process. Experimental results show that the proposed overall algorithm can save encoding time by 43.2% with negligible rate penalty of 0.47%. Compared with current state-of-art algorithms, the proposed algorithm can achieve the better trade-off between RD performance and complexity reduction. In the future, we will try hard to achieve more encoding time reduction while maintain the coding efficiency.

**Acknowledgements** This work is supported by the Natural Science Foundation of China (61771269, 61620106012, 61671258), Natural Science Foundation of Zhejiang Province (LY17F010005, LY16F010002) and Ningbo University Research Foundation (Science)/Discipline Project under Grant xkx11502. It is also sponsored by K.C. Wong Magna Fund in Ningbo University.

## References

1. Abdelrasoul M, Sayed MS, Goulart V (2017) Diagonal-based fast intra-mode decision algorithm for HEVC. *IET Image Process* 11(10):888–898
2. Belghith F, Kibeya H, Ayed M A B, Masmoudi N (2015) Fast coding unit partitioning method based on edge detection for HEVC intra-coding. *SIViP*, p 1–8



3. Bjontegaard G (2001) Calculation of average PSNR differences between RD-curves, ITU-T SG16 Q.6 document, VCEG-M33, Austin, US
4. Bossen F (2011) Common test conditions and software reference configurations, document JCTVC-F900, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), 6th meeting: Torino, IT
5. Chen ZY, Chang PC (2017) Rough mode cost-based fast intra coding for high-efficiency video coding. *J Vis Commun Image Represent* 43:77–88
6. Correa G, Assuncao PA, Agostini LV, da Sliva LA (2015) Fast HEVC encoding decisions using data mining. *IEEE Trans Circ Syst Video Technol* 25(4):660–673
7. Fang L, Li S, Kang X, Izatt J, Farsius S (2015) 3-D adaptive sparsity based image compression with applications to optical coherence tomography. *IEEE Trans Med Imaging* 34(6):1306–1320
8. Fang L, He N, Lin H (2017) CP tensor-based compression of hyperspectral images [J]. *JOSA A* 34(2):252–258
9. Fu W, Li S, Fang L, Benediktsson JA (2017) Adaptive spectral-spatial compression of hyperspectral image with sparse representation. *IEEE Trans Geosci Remote Sens* 55(2):671–682
10. Ismail Y, McNeely JB, Shaaban M, Mahmoud H, Bayoumi MA (2012) Fast motion estimation system using dynamic models for H.264/AVC video coding. *IEEE Trans Circ Syst Video Technol* 22(1):28–42
11. Jin YL, Park HW (2012) A fast mode decision method based on motion cost and intra prediction cost for H.264/AVC. *IEEE Trans Circ Syst Video Technol* 22(3):393–402
12. Jani L, Frank B, Woo-Jin H, Junghe M, Kemal U (2012) Intra coding of the HEVC standard. *IEEE Trans Circ Syst Video Technol* 12(22):1792–1802
13. Kim HS, Park RH (2016) Fast CU partitioning algorithm for HEVC using an online-learning-based bayesian decision rule [J]. *IEEE Trans Circ Syst Video Technol* 26(1):130–138
14. Kim KB, Jeon YG, Hong MC (2008) Variable step search fast motion estimation for H.264/AVC video coder. *IEEE Trans Consumer Electronics* 54(3):1281–1286
15. Lin W, Panusopone K, Baylon DM, Sun MT, Chen Z, Li H (2011) A fast sub-pixel motion estimation algorithm for H. 264/AVC video coding. *IEEE Trans Circ Syst Video Technol* 21(2):237–242
16. Lee J, Jeon B (2004) Fast mode decision for H.264. In Proc. ICME, p1131–1134
17. Li F, Jiao D, Shi G, Niu Y, Fan C, Xie X (2017) An AR based fast mode decision for H. 265/HEVC intra coding. *Multimed Tools and Appl*, p 1–19
18. Liao W, Yang D, Chen Z (2017) A fast mode decision algorithm for HEVC intra prediction. In Proc. ICIP, p 1–4
19. Liu Y, Nie L, Han L, Zhang L, Rosenblum DS (2015) Action2Activity: recognizing complex activities from sensor data. In Proc. IJCAI, p:1617–1623
20. Liu Y, Nie L, Liu L, Rosenblum SD (2016) From action to activity: sensor-based activity recognition. *Neurocomputing* 181:108–115
21. Liu Y, Zhang L, Nie L, Yan Y, Rosenblum D (2016) Fortune teller: predicting your career path. In Proc. AAAI, p 201–207
22. Liu L, Cheng L, Liu Y, Jia Y, Rosenblum D (2016) Recognizing complex activities by a probabilistic interval-based model. In Proc. AAAI, p:1266–1272
23. Liu Y, Zheng Y, Liang Y, Liu S, Rosenblum D (2016) Urban water quality prediction based on multi-task multi-view learning. *Proceedings of the international joint conference on artificial intelligence Google Scholar*
24. Liu X, Liu Y, Wang P, Lai CF, Chao HC (2017) An adaptive mode decision algorithm based on video texture characteristics for HEVC intra prediction. *IEEE Trans Circ Syst Video Technol* 27(8):1737–1748
25. Lu X, Xiao N, Hu Y, Martin G, Jin X, Wu Z (2016) A hierarchical fast coding unit depth decision algorithm for HEVC intra coding. In Proc VCIP, p 1–4
26. Min B, Cheung RC (2015) A fast CU size decision algorithm for the HEVC intra encoder. *IEEE Trans Circ Syst Video Technol* 25(5):892–896
27. Ismail Marzuki, Jonghyun Ma, Yong-Jo Ahn, Donggyu Sim (2016) A context-adaptive fast intra coding algorithm of high-efficiency video coding (HEVC). *J Real-Time Image Proc*, p 1–17
28. Ohm JR, Sullivan GJ, Schwarz H, Tan TK, Wiegand T (2012) Comparison of the coding efficiency of video coding standards—including high efficiency video coding (HEVC). *IEEE Trans Circ Syst Video Technol* 22(12):1669–1684
29. Preotjuc-Pietro D, Ye L, Hopkins D, Ungar L (2017) Beyond binary labels: political ideology prediction of twitter users. In Proc Meeting of the Association for Computational Linguistics, p 729–740
30. Ramezanpour M, Zargari F (2015) Early termination algorithm for CU size decision in HEVC intra coding. In Proc MVIP, p:45–48
31. Ruiz D, Fernández-Escribano G, Martínez JL, Cuenca P (2016) Fast intra mode decision algorithm based on texture orientation detection in HEVC. *Signal Process Image Commun* 44(C):12–28

32. Shang X, Wang G, Fan T, Li Y (2015) Fast CU size decision and PU mode decision algorithm in HEVC intra coding. In Proc ICIP, p 1593–1597
33. Shen L, Zhang Z, An P (2013) Fast CU size decision and mode decision algorithm for HEVC intra coding. *IEEE Trans Consum Electron* 59(1):207–213
34. Shen L, Zhang Z, Liu Z (2014) Effective CU size decision for HEVC intra coding. *IEEE Trans Image Processing* 23(10):4232–4241
35. Sullivan GJ, Ohm JR, Han WJ, Wiegand T (2012) Overview of the high efficiency video coding (HEVC) standard. *IEEE Trans Circ Syst Video Technol* 22(12):1871–1884
36. Wang JC, Wang JF, Yang JF, Chen JT (2007) A fast mode decision algorithm and its VLSI design for H.264/AVC intra-prediction. *IEEE Trans Circ Syst Video Technol* 17(10):1414–1422
37. Yang Y, Liu Q, Liu H, Yu L, Wang F (2015) Dense depth image synthesis via energy minimization for three-dimensional video. *Signal Process* 112:199–208
38. Yang Y, Wang X, Liu Q, Xu M, Yu L (2015) A bundled-optimization model of multiview dense depth map synthesis for dynamic scene reconstruction. *Inf Sci* 320:306–319
39. Zhang H, Ma Z (2014) Fast intra mode decision for high efficiency video coding (HEVC). *IEEE Trans Circ Syst Video Technol* 24(4):660–668
40. Zhang Y, Kwong S, Wang X, Yuan H, Pan ZQ, Xu L (2015) Machine learning-based coding unit depth decisions for flexible complexity allocation in high efficiency video coding. *IEEE Trans Image Process* 24(7):2225–2238
41. Zhang T, Sun MT, Zhao D, Gao W (2017) Fast intra-mode and CU size decision for HEVC [J]. *IEEE Trans Circ Syst Video Technol* 27(8):1714–1726
42. Zhang Y, Pan Z, Li N, Wang X, Jiang G, Kwong S (2017) Effective data driven coding unit size decision approaches for HEVC intra coding. *IEEE Trans Circ Syst Video Technol* PP(99):1–1
43. Zhao W, Onoye T, Song T (2015) Hierarchical structure-based fast mode decision for H.265/HEVC. *IEEE Trans Circ Syst Video Technol* 25(10):1651–1664
44. Zhu L, Zhang Y, Pan Z, Wang R, Kwong S, Peng Z (2017) Binary and multi-class learning based low complexity optimization for HEVC encoding. *IEEE Trans Broadcasting* 63(3):547–561



**Fen Chen** received her B.S. degree from Sichuan Normal College, China, in 1996, and M.S. degree from Institute of Optics and Electronics, Chinese Academy of Science, in 1999. She is now an associate professor in Faculty of Information Science and Engineering, Ningbo University, China. Her research interests mainly include digital signal processing, image processing and communications, multiview video processing.



**Defu Jin** received his B.S. degree in communication engineering from China Jiliang University in 2015. He is now studying as a postgraduate student in the Ningbo University, China. His research interests concentrate on image video compression.



**Zongju Peng** received his B.S. degree from Sichuan Normal College, China, in 1995, and M.S. degree from Sichuan University, China, in 1998, and received his Ph.D. degree from Institute of Computing Technology, Chinese Academy of Science in 2010. He is now a professor in Faculty of Information Science and Engineering, Ningbo University, China. His research interests mainly include image/video compression, multi-view video coding and video perception.



**Gangyi Jiang** received his M.S. degree from Hangzhou University in 1992, and received his Ph.D. degree from Ajou University, Korea, in 2000. He is now a professor in Faculty of Information Science and Engineering, Ningbo University, China. His research interests mainly include digital video compression and communications, multi-view video coding and image processing.



**Mei Yu** received her B.S. and M.S. degrees from Hangzhou Institute of Electronics Engineering, China, in 1990 and 1993, and Ph.D. degree from Ajou University, Korea, in 2000. She is now with the Faculty of Information Science and Engineering, Ningbo University, China. Her research interests include image/video coding and video perception.



**Hua Chen** received the M.S. degree and Ph.D. degree in Information and Communication Engineering from Tianjin University, Tianjin, China, in 2013 and 2017, respectively. From Aug. 2015 to Aug. 2016, he was a one-year visiting scholar at the Department of ECE, Concordia University. He is now with Faculty of Information Science and Engineering, Ningbo University, Ningbo, China. His research interests include array signal processing, MIMO radar and passive radar.