

# A fast inter-frame encoding scheme using the edge information and the spatiotemporal encoding parameters for HEVC

Zhiyao Yang<sup>1</sup> · Shuxu Guo<sup>1</sup> · Qinglong Shao<sup>1</sup>

Received: 18 May 2016 / Revised: 20 September 2016 / Accepted: 15 November 2016 /  
Published online: 28 November 2016  
© Springer Science+Business Media New York 2016

**Abstract** High Efficiency Video Coding (HEVC), as a novel video coding standard, has shown a better coding efficiency than all existing standards, such as H.264/AVC. It adopts a lot of new efficient coding tools, the most important one is the new hierarchical structures which include the coding unit (CU), prediction unit (PU) and transform unit (TU). However, the rate-distortion (RD) optimization process for all CUs, PUs, and TUs cause large computational costs. In this paper, a fast Inter-frame encoding scheme using the edge information and the spatiotemporal encoding parameters is proposed to reduce the encoder complexity of HEVC, which consists of a fast all  $2N \times 2N$  modes decision method, a fast CU depth level decision method and a fast PU mode decision method. This scheme uses edge information to express the structure complexity and uses the difference of edge information between current CU and its spatiotemporal CUs to express the edge similarity (ES) in one frame and the edge movement (EM) between two adjacent frames. And then, utilizes ES and EM as assistant parameters cooperate with CU depth levels and PU mode RD costs of spatiotemporal CUs to accomplish the early termination of CU split and the PU mode selection. The experimental results show that the proposed fast inter-frame encoding method can significantly reduce the computational costs with negligible RD loss. There are 53.7 and 54.9% encoding time savings on average, but only with average 1.5 and 1.8% Bjøntegaard difference bitrate (BDBR) losses for various test sequences under random access and low delay conditions, respectively.

**Keywords** High Efficiency Video Coding (HEVC) · Fast coding unit depth decision · Fast prediction unit mode decision · Edge information

---

✉ Zhiyao Yang  
naruto\_yang@163.com

<sup>1</sup> College of Electronic Science and Engineering, Jilin University, Changchun, China

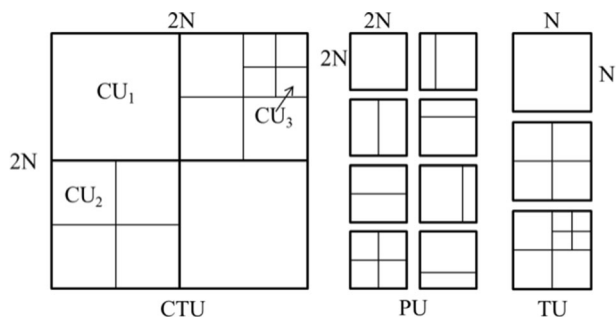
### 1 Introduction

High Efficiency Video Coding (HEVC) [17] is a new standardization for video compression which is proposed by the Joint Collaborative Team On Video Coding (JCT-VC) formed by the ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG). The goal of HEVC is to achieve 50% bit-rate reduction over H.264/AVC with the comparable video qualities.

Several new coding tools have been introduced in HEVC, the most important one is the new hierarchical structures which include the coding unit (CU), prediction unit (PU) and transform unit (TU). Figure 1 shows the new structures of CU, PU, and TU in HEVC. The basic coding unit in HEVC is the coding tree unit (CTU), which has a size selected by the encoder. The size of CTU can be chosen as  $64 \times 64$ ,  $32 \times 32$  or  $16 \times 16$ , as shown in Fig. 1, the CTU block (CTB) can be recursively split into four sub-CUs according to the CU depth, so the CU size is  $2N \times 2N$  ( $N = 32, 16, 8, 4$ ). CU is the root of PU and TU and the size of PU and TU cannot exceed the CU size. PU is a basic unit for motion estimation with symmetrical structures ( $2N \times 2N$ ,  $2N \times N$ ,  $N \times 2N$  and  $N \times N$  sizes, but  $N \times N$  is only for the smallest CU) or asymmetrical structures ( $2N \times nU$ ,  $2N \times nD$ ,  $nL \times 2N$  and  $nR \times 2N$  sizes). TU is the basic unit of transform which can be recursively split in a quad-tree structure and its size is  $N \times N$ . HEVC encoder takes advantage of rate-distortion optimization (RDO) to choose from those variable block sizes to specify the optimal CU size, PU coding mode and TU size. Although these structures improve the coding efficiency, the encoding complexity [3] also increases greatly. Therefore, it is of great importance to reduce the encoding complexity while maintaining the coding efficiency of HEVC.

In recent years, many methods have been proposed to reduce the complexity of HEVC encoder. Some of them have focused on the fast mode decision methods for HEVC intra-frame coding [9]. Zhang and Ma [21] proposed a fast intra mode decision for the HEVC encoder. There are two parts of their method, one is that they proposed a Hadamard cost based progressive rough mode search (pRMS) to accelerate the modes selection for intra prediction; the other one is that they compared the aggregated rate-distortion (RD) costs of the partial sub-CUs with the RD cost of the current CU to realize CU early split termination. Some researchers proposed fast TU decision methods. Kim [8] proposed a method that can determine the appropriate TU size effectively according to the position of the last non-zero transform coefficient and the number of zero transform coefficients. Shen et al. [15] used the Bayesian theorem [6, 7] to express the correlation between the TU size and the variance of residual coefficients to decrease the range of TU candidate modes. However, whether intra mode decision or TU size decision, only takes a little part of the HEVC complexity [3], so the

**Fig. 1** The typical example for CU, PU, and TU coding structures of HEVC



complexity reduction is relatively small, as their experimental results show that the encoding time reduction of [8, 15] is only 14–15% on average. Thus, many researches have focused on the fast Inter-frame encoding methods. References [11, 14] also utilize the Bayesian theorem to reduce HEVC complexity. Shen et al. [14] proposed an algorithm that collects relevant and computational-friendly features, such as motion vector (MV) magnitude, associate with the Bayesian decision rule to accelerate CU splitting. Lee et al. [11] proposed a fast CU size decision algorithm includes three approaches with SKIP mode decision, CU skip estimation and early CU termination. All the three approaches use the Bayes' rule and the statistical analysis to decide the RD cost thresholds and update process is performed to get those thresholds adaptively. Although those methods use Bayesian theorem can get a better complexity reduction, they need offline training to find thresholds which will need an extra time. Xiong et al. [18] proposed a fast inter CU decision method takes advantage of latent sum of absolute differences (SAD) and expresses the relationship between the motion compensation RD cost and the SAD cost as an exponential model. Xiong et al. [19] proposed another method based on Markov random field (MRF). The CU splitting can be modeled as a MRF inference problem and optimized by the Graphcut algorithm. As far as we know, using the spatiotemporal encoding parameters to early terminate CU splitting and PU mode decision is the most common method, such as references [1, 4, 10, 13, 16, 22, 23, 25, 26]. Lee et al. [10] proposed an efficient inter prediction mode decision method using spatial and upper-PU correlation to define the priority for each inter prediction mode and compare the priority with a threshold value which is acquired by offline experiments. Cen et al. [4] proposed a fast CU depth decision mechanism utilizes the distribution of CU depths in the same sequence frame to determine current CU depth range adaptively and skip a part of CU RD cost calculations at the current CU depth. Shen et al. [13] proposed a fast CU size decision algorithm which can determine CU depth range and skip some specific depth levels rarely used in the previous frame and neighboring CUs. Ahn et al. [1] proposed a fast CU encoding scheme utilizes sample-adaptive-offset (SAO) parameters as the spatial parameter to estimate the texture complexity and the motion vectors, TU size and coded block flag information as the temporal parameters to estimate the temporal complexity. Zhang et al. [22] analyzed the correlation of the current CTU and spatial-temporal neighboring CTUs to reduce the depth search range. Zhou et al. [26] also analyzed the correlation of spatiotemporal CUs and found the depth monotonicity between current CU and its neighboring CUs. Shen et al. [16] proposed a fast inter-mode decision algorithm jointly using the spatiotemporal correlation of CUs and the inter-level correlation of quadtree structure to statistically analyze the coding information among the adjacent CUs and the prediction mode distribution at each depth level. Zhang et al. [23] utilized the spatiotemporal correlation of CUs to divide current CU into three mode regions with different range of CU depth levels. Zhong et al. [25] analyzed the occurrence probability of CU segmentation and partition mode in two adjacent frames and utilized the spatial correlation of CU size and partition mode between the corresponding CU and its four surrounding CUs to save computational complexity. Zhao et al. [24] first proposed a fast inter-mode decision algorithm uses the depth information of CUs in previous frame, and then cooperate it with a hardware-oriented fast intra-mode decision algorithm. In addition to use the spatiotemporal correlation, there are some special methods [5, 20]. Xiong et al. [20] presented a fast CU selection algorithm based on pyramid motion divergence (PMD) features which are calculated with estimated optical flow of the down sampled frames. Correa et al. [5] proposed schemes based on decision trees obtained through data mining technique which is belong to machine learning technology. Lee et al. [12] proposed an early skip mode decision with emphasis on coding quality.

However, these algorithms did not sufficiently discover the video contents features which can directly reflect the image complexity in one frame and contents similarity between adjacent frames. In this paper, we propose a fast Inter-frame encoding scheme jointly using the edge information and the spatiotemporal encoding parameters to reduce the encoder complexity of HEVC. It has three approaches, a fast all  $2N \times 2N$  modes decision, a fast CU depth level decision and a fast PU mode decision. This scheme first calculates the mean value of the edge information (include the gradient component and the phase component) in one CU. The mean gradient value of one CU is used to express the complexity degree of CU. And then, calculates the difference of mean gradient value and mean phase value between current CU and its spatiotemporal CUs to expression the edge similarity (ES) in the same frame and the edge movement (EM) between two adjacent frames. Second, use the mean gradient value to realize the fast all  $2N \times 2N$  modes decision and utilizes ES and EM as reference parameters cooperate with CU depth levels and PU RD costs of spatiotemporal CUs to accomplish the early termination of CU split and the PU mode selection, respectively.

The rest of this paper is organized as follows. Section 2 observes the relationship between image edge distribution and different CU sizes or PU modes and then analysis the mean gradient value of one CU, ES and EM how to be used in fast all  $2N \times 2N$  modes decision, CU splitting and PU mode decision. The proposed fast encoding scheme is detailed in Section 3. Section 4 gives the experimental results and performance analyses. Section 5 concludes our work in this paper.

## 2 Observation and analysis of the relationship between edge information and CU structures

There are some researches [1, 13] have shown that the CU depth and PU mode are related to the video contents. If a CU located in the homogenous region of a video sequence, it always has a high probability to be encoded by the large size block or large PU mode. On the contrary the CU in complex region is often encoded as a small block or small PU mode. On the other hand, edge is one basic characteristic of picture and it can reflect the position of an object and the texture complexity. Edge detection is a kind of technique in image processing. The edge distribution information can be obtained through edge detection. There are many algorithms for edge detection, in our work, we need to decrease the encoding complexity so there should not introduce extra computation complexity into our scheme. Therefore, we selected some frequently-used edge detection operators (such as Sobel operator, Canny operator, Prewitt operator and so on) to get the edge information and make a contrast of those operators, and finally choose the Prewitt operator which is less time consumption and easy to realize the detection of edge to get the edge information of each frame in video sequences and the operator is shown in Fig. 2.

**Fig. 2** The template of Prewitt operator

1	1	1
-1	-1	-1

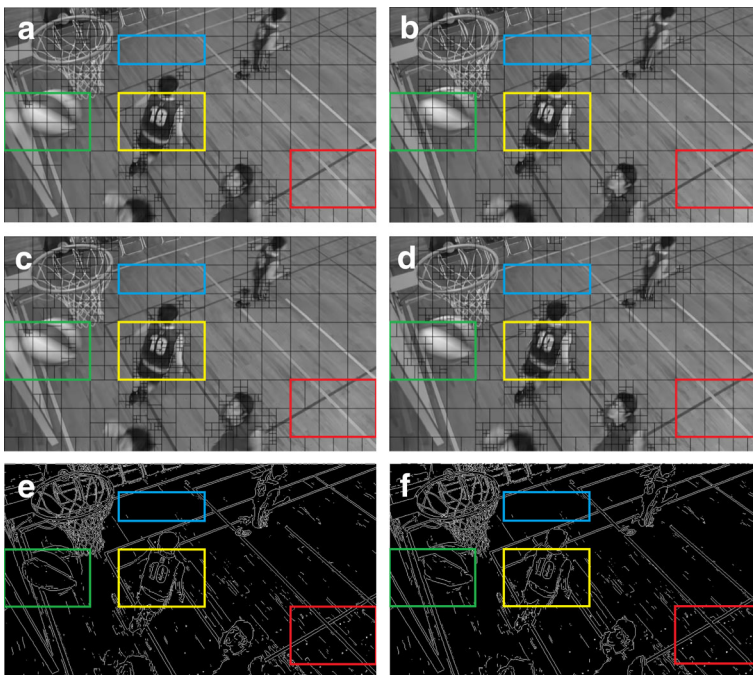
-1		1
-1		1
-1		1

The edge value is calculated by formula (1).

$$\left\{ \begin{array}{l} G_x = \left| p(i-1, j+1) + p(i, j+1) + p(i+1, j+1) - p(i-1, j-1) - p(i, j-1) - p(i+1, j-1) \right| \\ G_y = \left| p(i-1, j-1) + p(i-1, j) + p(i-1, j+1) - p(i+1, j-1) - p(i+1, j) - p(i+1, j+1) \right| \\ E(i, j) = G_x + G_y \\ Ph(i, j) = \arctan(G_y/G_x) \end{array} \right. \quad (1)$$

In (1),  $i$  and  $j$  are the coordinate of pixel. The  $p(i, j)$  is the pixel at position  $(i, j)$  in one CU.  $G_x$  is the horizontal direction edge component and  $G_y$  is vertical direction edge component.  $E(i, j)$  is the edge gradient value at position  $(i, j)$  in one CU.  $Ph(i, j)$  is the edge phase value at position  $(i, j)$  in one CU. In order to directly display the edge distribution situation of video sequences, we choose frame 2 and 3 of test sequence “BasketballDrill” encoded at QP = 32 as an example. The distribution of CU depth levels, PU modes and edge information are shown in Fig. 3.

Test sequence “BasketballDrill” is recommended by JCT-VC and its resolution is  $832 \times 480$  and the frame rate is 50fps. Figure 3a and b demonstrate the CU depth levels distribution in the luminance image of two adjacent frames 2 and 3, respectively. Figure 3c and d demonstrate the PU modes distribution in these two adjacent frames, respectively and Fig. 3e and f are the corresponding edge information distribution of these two frames,



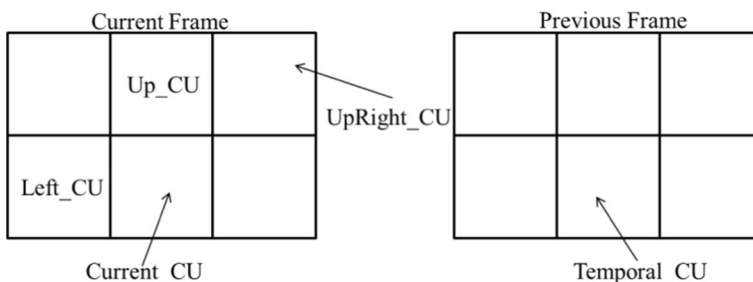
**Fig. 3** The distribution of CU depth levels, PU modes and edge information in two adjacent frames (a) CU depth levels distribution of frame 2 (b) CU depth levels distribution of frame 3 (c) PU modes distribution of frame 2 (d) PU modes distribution of frame 3 (e) Edge information distribution of frame 2 (f) Edge information distribution of frame 3

respectively. By comparing yellow block part in Fig. 3a, c and e or b, d and f) the conclusion is that the block with more edges will be easy to be encoded in a high depth (small CU size) and small PU mode. On the contrary, through the comparison of blue block, those CUs with fewer edges will be encoded in a low depth (large CU size) and large PU mode. So our first idea is use the mean edge gradient value in each CU to represent current CU's complexity degree.

Meanwhile, as shown in Fig. 3a and b, c and d those spatially adjacent CUs usually have similar CU depth levels and PU modes, such as the CUs in blue and yellow block, respectively. In addition to this, a lot of CUs in frame 3 have the same depth levels and PU modes distribution as corresponding position CUs in frame 2. Therefore, the optimal depth level and PU mode of the current CU may have a strong correlation with its spatially adjacent CUs and corresponding temporal CU. In references [1, 13, 16, 22, 26], there have the same conclusion, but most of them use this relationship through statistics analysis or the offline training. In this paper, the edge information is introduced to estimate the similarity between current CU and its spatially adjacent CUs or corresponding temporal CU. Then, according to the edge similarity we can find the best fit CU from the spatiotemporal adjacent CUs, and the coding information (such as CU depth level and PU RD cost) of those best fit CUs can be used for current CU, which will save some useless computational process and it will speed up the coding process.

As shown in the green block of Fig. 3a and b, c and d there exist an object moving between two adjacent frames, so the depth levels and PU modes of those CUs in this green block are changing. On the contrary, in the red block of Fig. 3a and b, c and (d), the depth levels and PU modes almost have no change and there are no objects moving. By comparing green and red block parts in Fig. 3e and f, we found that the objects moving and the similarity of CUs can be represented by difference of edge information. So the second idea of our works is using the difference of edge information to reflect the object similarity and movement in each CU, and then we cooperates that information with CU depth level and PU RD cost of spatiotemporal adjacent CUs to fulfill the fast CU depth level decision and the fast PU mode decision.

According to references [1, 13, 16, 22, 26], we choose four frequently-used spatiotemporal CUs as predictors. They are the corresponding position temporal CU ( $T\_CU$ ), the left CU ( $L\_CU$ ), the up CU ( $U\_CU$ ) and the upright CU ( $Ur\_CU$ ) of current CU ( $C\_CU$ ), as shown in Fig. 4. Combine our two ideas, the mean edge gradient value and mean edge phase value of current and its predictors are expressed as  $mE(C\_CU)$  and  $mPh(C\_CU)$ ,  $mE(T\_CU)$  and  $mPh(T\_CU)$ ,  $mE(L\_CU)$  and  $mPh(L\_CU)$ ,  $mE(U\_CU)$  and  $mPh(U\_CU)$ ,  $mE(Ur\_CU)$  and  $mPh(Ur\_CU)$  which can be calculated as follow.



**Fig. 4** Spatial and temporal correlations of CUs

$$\begin{cases} mE(X) = 1/(2N \times 2N) \sum_{i=1}^{2N} \sum_{j=1}^{2N} E(i, j) \\ mPh(X) = 1/(2N \times 2N) \sum_{i=1}^{2N} \sum_{j=1}^{2N} Ph(i, j) \end{cases} \quad (2)$$

Where  $X \in \{C\_CU, T\_CU, L\_CU, U\_CU, Ur\_CU\}$ . Then, the similarity of CUs in one frame and the movement of CUs in two adjacent frames which are respectively represent by  $ES(Y)$  and  $PhS(Y)$  ( $Y \in \{L\_CU, U\_CU, Ur\_CU\}$ ),  $EM$  and  $PhM$  can be calculated by formula (3).

$$\begin{cases} ES(Y) = mE(C\_CU) - mE(Y) \\ PhS(Y) = mPh(C\_CU) - mPh(Y) \\ EM = mE(C\_CU) - mE(T\_CU) \\ PhM = mPh(C\_CU) - mPh(T\_CU) \end{cases} \quad (3)$$

Through a large number of experiments, we give a condition that is  $ES(Y) = 0$  and  $PhS(Y) \in [-\pi/64, \pi/64]$  which means that the mean edge gradient value of current CU is same to its spatial adjacent CUs, and the mean edge phase value of current CU is close to its spatial adjacent CUs, so that the edge distribution of current CU maybe similar to its spatial adjacent CUs. On the other hand, when  $EM = 0$  and  $PhM \in [-\pi/64, \pi/64]$  the mean edge gradient value of current CU is same to its co-located CU, and the mean edge phase value of current CU is close to its co-located CU in previous frame, so that the edge distribution of current CU maybe similar to its co-located CU in previous frame which means that there exist no object moving in current CU. Therefore, if  $ES(Y) = 0$  and  $PhS(Y) \in [-\pi/64, \pi/64]$  or  $EM = 0$  and  $PhM \in [-\pi/64, \pi/64]$  the coding information of those spatiotemporal CUs may be used in current CU directly.

### 3 Proposed fast inter-frame encoding scheme for HEVC encoding

As described in Section 2, there are two ideas have been proposed one is using the mean edge gradient value to reflect the CU structure complexity, the other one is combining the depth levels and the RD costs of spatiotemporal adjacent CUs with ES and PhS, EM and PhM to realize our fast CU depth level decision and fast PU mode decision. In this part, we will describe our fast inter-frame encoding scheme from three aspects in detail.

#### 3.1 Fast all $2N \times 2N$ modes decision using the mean edge gradient value

In our work we focus on inter-frame prediction and through the analysis of HEVC reference software HM we divide PU mode into 4 parts, they are Skip mode,  $2N \times 2N$  mode,  $2N \times N$  ( $N \times 2N$ ) mode, asymmetrical and other modes. Skip mode is a special mode in HEVC encoder, the CU size of Skip mode is also  $2N \times 2N$  and when current CU choose Skip mode as its best prediction mode it will use the adjacent CUs coding information directly, so Skip mode is the simplest one among those prediction mode. In our fast all  $2N \times 2N$  modes decision method we combine Skip mode and  $2N \times 2N$  mode collectively call “all  $2N \times 2N$  modes”. As we analyzed in Section 2, when there is a simple edge distribution situation in one CU means that edge value of this CU is small and the prediction mode will be simple, on the contrary when the edge distribution situation is complex in one CU means that edge value of this CU is large and the prediction mode will be complex. So we choose the mean edge gradient value of current CU as a parameter for fast all  $2N \times 2N$  modes decision method.

In our fast all  $2N \times 2N$  modes decision method, when mean edge gradient value of current CU is smaller than a threshold, current CU choose all  $2N \times 2N$  modes as its best prediction mode, and the other prediction modes will be skipped. In [16], there is a conclusion has been proved that when the depth level of current CU is small the probability of choosing all  $2N \times 2N$  modes as the best mode will be low. In our work, the max CU depth is 4 (depth level can be 0, 1, 2, 3). So, in order to improve the accuracy rate of our method the fast all  $2N \times 2N$  modes decision method only be used when CU depth level larger than 0.

As for the threshold, through a large number of experiments, the threshold has been set as the mean edge gradient value of the whole frame which can be represent by “ $mfE$ ”. If “ $mE(X) < mfE$ ” and the depth level of current CU is larger than 0, the prediction of current CU is set as all  $2N \times 2N$  modes and the  $2N \times N$  ( $N \times 2N$ ) mode, asymmetrical and other modes will be skipped. For simplify the expression, we set the depth level of current CU is larger than 0 and  $mE(X) < mfE$  as “Condition 1”. In order to prove the accuracy of the fast all  $2N \times 2N$  modes decision method we gave the all  $2N \times 2N$  modes hit rates of our method. We choose four test sequences “BasketballPass” ( $416 \times 240$ ), “BQMall” ( $832 \times 480$ ), “Cactus” ( $1920 \times 1080$ ), and “Traffic” ( $2560 \times 1600$ ) and each of them was encoded at QP = 22, 27, 32, and 37 using HM 16.0. The result is shown in Table 1.

As shown in Table 1, the hit rates of those test sequences are all exceed 90%, it means that our proposed fast all  $2N \times 2N$  modes decision method have a high accuracy.

### 3.2 Fast CU depth level decision using depth level information with ES and PhS, EM and PhM

As we know there is a great similarity about the depth levels distribution between current CU and its spatial adjacent CUs or co-located CU in previous frame, just like those CUs we have chosen in Section 2. So that, the depth levels of those CUs can be used to predict current CU’s depth level. However, there are a lot of objects in one frame, in different objects the contents in CU will have a great difference and if we use these CUs to predict current CU it will bring great errors. On the other hand, in two adjacent frames the objects may have a movement. It also can bring a great difference between current CU and its co-located CU in previous frame, in other words, it also will make a predict error. In order to reduce those predict errors, we proposed to use ES and PhS, EM and PhM to respectively estimate the objects similarity and movement between current CU and its spatiotemporal adjacent CUs in advance. When  $ES(Y) = 0$  and  $PhS(Y) \in [-\pi/64, \pi/64]$  or  $EM = 0$  and  $PhM \in [-\pi/64, \pi/64]$ , the corresponding CUs may have a great similarity, so the depth levels can be used directly as candidates. For simplify the expression, we set  $ES(Y) = 0$  and  $PhS(Y) \in [-\pi/64, \pi/64]$  or  $EM = 0$  and  $PhM \in [-\pi/64, \pi/64]$  as “Condition 2”.

In our work, we use  $D(C)$  and  $D(Z)$  ( $Z \in \{T\_CU, L\_CU, U\_CU, Ur\_CU\}$ ) to represent the depth level of current CU and its spatiotemporal adjacent CUs, respectively. To accelerate current CU splitting, those candidate depths are set as the max depth level for current CU. It means that when  $D(C) > D(Z)$  the splitting process will stop. In order to confirm the effectiveness of this method, we give the hit rates of this method only use the depth level

**Table 1** The hit rates of all  $2N \times 2N$  modes when Condition 1 is true

Sequences	BasketballPass	BQMall	Cactus	Traffic	Average
Hit rates	91.5%	90.9%	90.4%	93.1%	91.5%



and the method use the depth level with “Condition 2”. We used the same test sequences and same test conditions in Section 3.1 and the result is shown in Table 2.

The  $P(D(C) \leq n | D(Z) = n)$  ( $n = 0, 1, 2$ ) in Table 2 means the probability when candidate depth level is  $n$ , current CU’s depth level is smaller than  $n$ . From Table 2 we can see that when only used the candidate depth level in spatiotemporal adjacent CUs, most of the sequence have a hit rate higher than 50%, it verified the similarity of depth levels between current CU and its spatiotemporal neighbors. On the other hand, when we add “Condition 2” to the depth level only method, the hit rate promote greatly from 60 to 90%. Thus, the using of edge similarity and edge movement can greatly reduce the predict errors which created by depth level only predict method. But when  $D(Z) = 3$ , this predict method is equal to original HEVC CU splitting process. Therefore, we proposed a special method for  $D(Z) = 3$ .

Through the statistical analysis we found that when  $D(Z) = 3$  the  $D(C)$  always be larger than 0. So this fast CU depth level decision method for depth 3 is that when  $D(Z) = 3$  current CU’s depth level 0 will be skipped. However, ES and PhS, EM and PhM cannot be used in this method as above, because the more differences between current CU and its spatiotemporal adjacent CUs the higher possibility current CU to be encoder in high depth. On the other hand, in one frame, the object in current CU is always different from its adjacent CUs it means that the ES and PhS cannot reflect the object changing precisely, so ES and PhS cannot be used in this special method. Thus, we gave a new restricted condition that is the  $EM \neq 0$  and  $PhM \in [-\pi/2, -\pi/64) \cup (\pi/64, \pi/2]$  which will reflect the movement of object in current CU. For simplify the expression, we set  $EM \neq 0$  and  $PhM \in [-\pi/2, -\pi/64) \cup (\pi/64, \pi/2]$  as “Condition 3”. Here, the hit rates of this special method is shown in Table 3. The same test sequences and test conditions are adopted. From Table 3, we can see that when  $D(Z) = 3$  current CU rarely choose depth 0 as the best CU size. The hit rate of depth only method is average 77.6%. It verified our discovery that when  $D(Z) = 3$  the  $D(C)$  always be larger than 0.

When “Condition 3” was jointly used with depth only predicts method, the hit rate have a progress from average 77.6 to 90.9% which can make this method more accurate. So the fast CU depth level decision method can be briefly described as follow:

- Step 1 When “Condition 3” is true, if  $D(Z) = 3$  current CU will skip the detection of depth level 0, go to next depth level detection, else depth level 0 will be calculated and then go to Step 2;
- Step 2 When “Condition 2” is true, if  $D(Z) = n$  ( $n = 0, 1, 2$ ) the max depth level of current CU is set as  $n$ , only calculate the depth level smaller than  $n$  or equal to  $n$ , else do the original HEVC method.

**Table 2** The hit rates comparison of the candidate depth level only method and the method corporate candidate depth level with “Condition 2” when candidate depth smaller than 3

Sequences		BasketballPass	BQMall	Cactus	Traffic	Average
Using candidate depth level with “Condition 2”	$P(D(C) \leq 0   D(Z) = 0)$	95.8%	93.8%	93.0%	92.1%	93.7%
	$P(D(C) \leq 1   D(Z) = 1)$	95.6%	93.8%	89.0%	90.1%	92.1%
	$P(D(C) \leq 2   D(Z) = 2)$	95.1%	90.8%	90.2%	90.3%	91.6%
Only use candidate depth level	$P(D(C) \leq 0   D(Z) = 0)$	41.0%	53.5%	65.0%	70.9%	57.6%
	$P(D(C) \leq 1   D(Z) = 1)$	52.4%	54.5%	60.8%	68.5%	59.1%
	$P(D(C) \leq 2   D(Z) = 2)$	77.4%	55.0%	68.5%	68.5%	67.4%

**Table 3** The hit rates comparison of the candidate depth level only method and the method corporate candidate depth level with “Condition 3” when candidate depth is 3

Sequences	P(D(C) > 0 D(Z) = 3) with “Condition 3”	P(D(C) > 0 D(Z) = 3) Only depth
BasketballPass	92.1%	76.5%
BQMall	92.0%	82.5%
Cactus	90.9%	81.2%
Traffic	88.5%	70.1%
Average	90.9%	77.6%

### 3.3 Fast PU mode decision method using RD cost with EM and PhM

In [1, 13] the RD cost of spatiotemporal adjacent CUs have been used to realize the fast PU mode decision. Just like Section 3.2, ES and PhS, EM and PhM can also be used with RD cost to choose a reasonable threshold for fast PU mode decision. As described in Section 3.1, we divided those PU mode into four modes and in our work the RD costs of those modes are represented by R (U) (U ∈ {Skip, 2 N × 2 N, 2 N × N (N × 2 N), asymmetrical and others}). Meanwhile, the best mode RD costs of spatiotemporal adjacent CUs are represented by Rb (V) (V ∈ {T\_CU, L\_CU, U\_CU, Ur\_CU}). Different from depth level method in part 3.2, we only use EM and PhM to restrict mode detection. Because RD cost is directly calculated by inter-frame prediction and EM and PhM can directly reflect the relationship between current CU and it temporal adjacent frame, but ES and PhS only represent this relationship in one frame. In our work, EM and PhM is calculated for every CUs in each frame, so the spatial adjacent CUs of current CU have its own EM and PhM value, here we use EM (K) and PhM(K) (K ∈ {L\_CU, U\_CU, Ur\_CU}) to represented them. When  $EM = 0$  and  $PhM \in [-\pi/64, \pi/64]$  or  $EM(K) = 0$  and  $PhM(K) \in [-\pi/64, \pi/64]$ , there may have no objects moving in current CU or the spatial adjacent CUs which means that the RD cost of co-located CU or the spatial adjacent CUs may be similar to current CU so that the corresponding Rb (T\_CU) or Rb (K) can be chosen as a candidate. For simplify the expression, we set  $EM = 0$  and  $PhM \in [-\pi/64, \pi/64]$  or  $EM(K) = 0$  and  $PhM(K) \in [-\pi/64, \pi/64]$  as “Condition 4” and the reverse condition of “Condition 4” as “Condition 5”. Then the reasonable threshold is selected from those candidates by formula (4).

$$Th = \begin{cases} \max\{Rb(T\_CU), Rb(K)\} & \text{if “Condition 4” is true} \\ \min\{Rb(T\_CU), Rb(K)\} & \text{if “Condition 5” is true} \end{cases} \tag{4}$$

The  $Th$  in formula (4) is the threshold for fast PU mode decision. However, in [16] we know that the Skip mode occupied a large part of optimal PU mode. In order to improve the accuracy of our proposed method we gave a special threshold for Skip mode as follow.

$$Ths = \begin{cases} \max\{Rb(T\_CU), Rb(K)\} & \text{if “Condition 4” is true} \\ 1/M \sum_{i=0}^M Rs(i) & \text{if “Condition 5” is true} \end{cases} \tag{5}$$

In (5),  $Ths$  is the threshold for fast Skip mode decision,  $Rs(i)$  is the RD cost of each CU in previous frame which choose Skip mode as its best PU mode and  $M$  is the number of those CUs. It means that when “Condition 5” is true the mean RD cost value of those CUs which choose Skip mode as its best mode in previous frame will be chosen as threshold for Skip mode. In addition, in different depth level current CU will have a RD cost which is acquired

from the quarter part of this CU and will approximately be a quarter of the RD cost in previous depth level. So we divided  $Th$  and  $Ths$  to fit each depth level as follow.

$$Th(\alpha) = \frac{1}{4^\alpha} Th, \alpha = 0, 1, 2, 3 \tag{6}$$

$$Ths(\alpha) = \begin{cases} \frac{1}{4^\alpha} Ths \text{ if "Condition 4" is true} \\ 1/M_\alpha \sum_{i=0}^{M_\alpha} Rs_\alpha(i) \text{ if "Condition 5" is true} \end{cases}, \alpha = 0, 1, 2, 3 \tag{7}$$

$Th(\alpha)$  and  $Ths(\alpha)$  are the thresholds for current CU in depth level “ $\alpha$ ”,  $Rs_\alpha(i)$  is the RD cost of CU in depth level “ $\alpha$ ” in previous frame which choose Skip mode as its best mode and  $M_\alpha$  is the number of those CUs. In each CU the threshold will be calculated and then compare it with the RD cost of current CU in each depth level in this order like Skip mode,  $2 N \times 2 N$  mode,  $2 N \times N$  ( $N \times 2 N$ ) mode, asymmetrical and other modes. If  $R(U) < Th(\alpha)$  ( $R(\text{Skip}) < Ths(\alpha)$ ) the mode detection will stop. If “U” is asymmetrical and other modes the process of mode decision is same to original HEVC mode decision method. To prove the effectiveness of this method we use the same test sequences and same test conditions in Section 3.1 to count the hit rates of mode selection which is shown in Table 4.

From Table 4 we can see the hit rates of the proposed fast PU mode decision method is reach up to 93%. It verified that the proposed method will have a high accuracy. So the fast PU mode decision method is that compare the calculated threshold with the RD cost of current PU mode if RD cost is smaller than threshold the rest of mode will be skipped.

Those two methods used the previous frame to predict current CU depth level and PU mode that will bring error accumulation so the basic reference frame will be updated by original HEVC encoding method at the beginning of group of picture (GOP).

### 3.4 Overall algorithms

In order to express our proposed method clearly, first we list the restricted condition as follow.

- Condition 1 The depth level of current CU is larger than 0 and  $mE(X) < mfE$ .
- Condition 2  $ES(Y) = 0$  and  $PhS(Y) \in [-\pi/64, \pi/64]$  or  $EM = 0$  and  $PhM \in [-\pi/64, \pi/64]$ .
- Condition 3  $EM \neq 0$  and  $PhM \in [-\pi/2, -\pi/64) \cup (\pi/64, \pi/2]$ .
- Condition 4  $EM = 0$  and  $PhM \in [-\pi/64, \pi/64]$  or  $EM(K) = 0$  and  $PhM(K) \in [-\pi/64, \pi/64]$ .
- Condition 5 reverse condition of “Condition 4”.

And then, the proposed overall fast inter-frame encoding method algorithm is described as follow and the flowchart of the overall fast inter-frame encoding scheme is shown in Fig. 5.

**Table 4** The hit rates of using the given threshold for skip,  $2 N \times 2 N$  and  $2 N \times N$  ( $N \times 2 N$ ) mode

Sequences		BasketballPass	BQMall	Cactus	Traffic	Average
The hit rates of proposed fast PU mode decision	P(R (Skip) < Th( $\alpha$ ) or Ths( $\alpha$ ))	93.6%	92.5%	94.3%	95.2%	93.9%
	P(R ( $2 N \times 2 N$ ) < Th( $\alpha$ ))	92.2%	91.7%	93.9%	95.7%	93.4%
	P(R ( $2 N \times N$ or $N \times 2 N$ ) < Th( $\alpha$ ))	92.8%	91.5%	93.6%	94.9%	93.2%

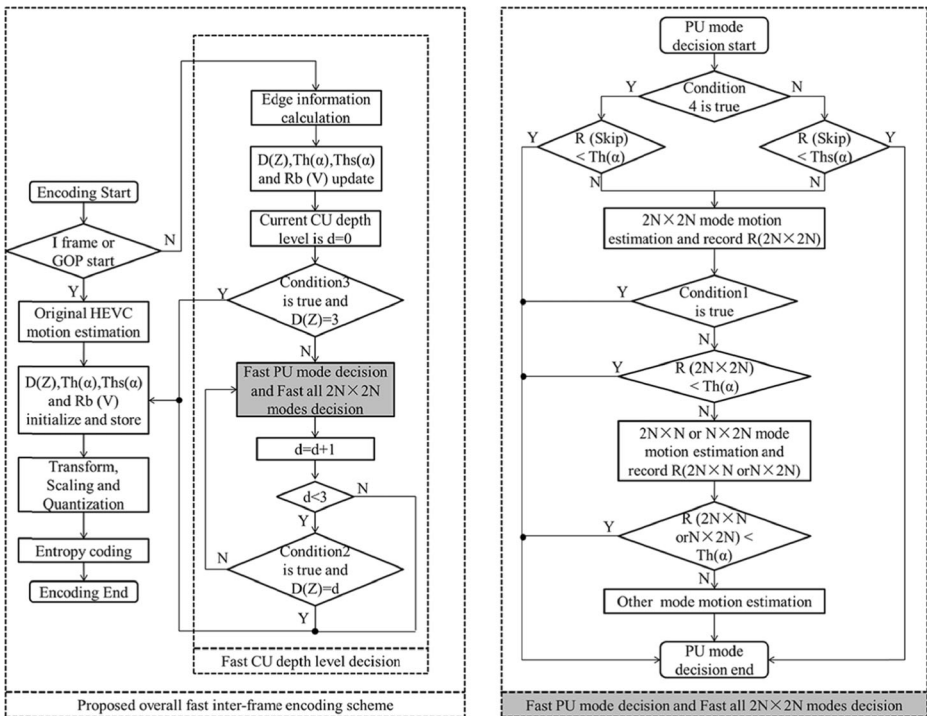


Fig. 5 Flowchart of the overall fast inter-frame encoding scheme

- Step 1 Start encoding, if current frame is I frame or GOP beginning go to Step 10, otherwise calculate the edge information include mean edge gradient value and mean edge phase value of current CU and its spatiotemporal adjacent CUs to determine whether Condition 1 to 5 is fulfilled or not and then go to Step 2.
- Step 2 Get the encoding information include  $D(Z), Th(\alpha), Ths(\alpha)$  from spatiotemporal adjacent CUs. Current CU depth level is “d” and  $d=0$ , if Condition 3 is true and  $D(Z)=3$  go to Step 9, otherwise go to Step 5.
- Step 3 Current CU depth level plus one. If current depth level is large than 3 go to Step 9, otherwise go to Step 4.
- Step 4 If Condition 2 is true and  $D(Z)$  is equal to current depth level then go to Step 9, otherwise go to Step 5.
- Step 5 PU mode decision start, the detection order of those prediction mode is that Skip mode is first detected, second is  $2N \times 2N$  mode, third is  $2N \times N$  (or  $N \times 2N$ ) mode and the last one is asymmetrical mode. For Skip mode if Condition 4 is true the threshold is set as  $Th(\alpha)$  and if Condition 4 is false (Condition 5 is true) the threshold is set as  $Ths(\alpha)$ . For other mode the threshold is set as  $Th(\alpha)$ . If RD cost of Skip mode is smaller than the threshold Skip mode will be the best prediction mode and then go to Step 3, otherwise go to Step 6
- Step 6 For  $2N \times 2N$  mode the fast all  $2N \times 2N$  modes is used first if Condition 1 or RD cost of  $2N \times 2N$  mode is smaller than  $Th(\alpha)$   $2N \times 2N$  mode will be the best prediction mode and then go to Step 3, otherwise go to Step 7.

- Step 7 For  $2N \times N$  (or  $N \times 2N$ ) mode, if RD cost of  $2N \times N$  (or  $N \times 2N$ ) mode is smaller than  $T_h(\alpha) 2N \times N$  (or  $N \times 2N$ ) mode will be the best prediction mode and then go to Step 3, otherwise go to Step 8.
- Step 8 Detected the asymmetrical mode and then go to Step 3.
- Step 9 Update the encoding information of spatiotemporal adjacent CUs, and then execute the other encoding processes such as transform, scaling, quantization and Entropy, if current frame is the last frame of test sequence the encoding process will be stop, otherwise read next frame and go to Step 1.
- Step 10 Do original HEVC encoding method and store the encoding information of spatiotemporal adjacent CUs, and then execute the other encoding processes such transform, scaling, quantization and Entropy, if current frame is the last frame of test sequence the encoding process will be stop, otherwise read next frame and go to Step 1.

## 4 Experimental results

In order to evaluate the effectiveness of the proposed fast inter-frame encoding scheme, we implemented it into the recent HEVC reference software HM16.0 and common test conditions are shown in [2].

There are 8 test conditions (reflect combinations of 8 and 10 bit coding, and of intra-only, random-access, and low-delay settings) defined in this standardization, because of the proposed method is used for inter-frame, so we choose two cases random access (RA) and low-delay B (LD) to conduct our experiments. The CTU size is set as  $64 \times 64$  and max CU depth is 4. Fast search mode and fast encoder decision are enabled and the search range set as 64. The test sequences we used are recommended by JCT-VC with five classes in different resolutions (class A  $2560 \times 1600$ / class B  $1920 \times 1080$ / class C  $832 \times 480$ /class D  $416 \times 240$ / class E  $1280 \times 720$ ) and four QPs of 22, 27, 32, and 37 are selected to evaluate the proposed algorithm. The performance of the proposed scheme is measured by the Bjøntegaard difference bitrate (BDBR) and encoding time saving. The variation of encoding time  $\Delta T$  is calculated as follow.

$$\Delta T = \frac{T_{HEVC} - T_{proposed}}{T_{HEVC}} \times 100\% \quad (8)$$

$T_{HEVC}$  and  $T_{proposed}$  are the encoding time of using the original algorithm in HEVC reference software and the proposed method in our work, respectively. Table 5 shows the experimental results for our proposed fast CU encoding scheme which include the fast all  $2N \times 2N$  modes decision, fast CU depth level decision and the fast PU mode decision. As shown in Table 5, the fast all  $2N \times 2N$  modes decision can achieve average 24.7% encoding time reduction and only with average 0.7% BDBR loss, the fast CU depth level decision method can achieve average 24.9% encoding time reduction and only with average 0.4% BDBR loss and the fast PU mode decision method can achieve average 34.7% encoding time reduction and only with average 0.6% BDBR loss. From these results we can see that our proposed three methods can achieve good complexity reductions and with little BDBR losses. The combination of these three methods can achieve average 56.9% encoding time reduction and only with average 1.5% BDBR loss. For the test sequences in Class E, the encoding time reductions exceed 70% and most of the high resolution sequences, such as the test sequences in Class A and B, the time reductions exceed 60%, it means that our method will be useful in high resolution video.

**Table 5** Performance of proposed fast inter-frame encoding scheme under RA condition

Class resolution	Sequences	Proposed fast all 2 N × 2 N modes decision		Proposed fast CU depth level decision		Proposed fast PU mode decision		Proposed overall fast encoding scheme	
		BDBR(%)	ΔT(%)	BDBR(%)	ΔT(%)	BDBR(%)	ΔT(%)	BDBR(%)	ΔT(%)
Class A 2560 × 1600	Traffic	0.8	28.3	0.5	26.8	0.9	44.5	2.0	63.2
	PeopleOnStreet	1.1	25.5	0.2	12.1	0.7	25.0	1.5	48.4
	Kimono1	0.8	36.2	0.5	10.3	1.1	32.7	2.5	58.6
Class B 1920 × 1080	ParkScene	1.0	28.2	0.2	21.8	0.9	42.7	1.7	61.8
	Cactus	0.7	30.0	0.2	21.1	0.7	35.7	1.6	60.3
	BasketballDrive	0.6	31.9	0.6	12.2	0.8	32.5	2.1	57.5
Class C 832 × 480	BQTerrace	0.5	28.8	0.2	20.7	0.7	41.1	1.5	60.0
	BasketballDrill	0.7	21.5	1.2	29.5	0.3	29.3	1.5	51.1
	BQMall	1.0	18.6	0.7	27.3	0.5	29.1	1.4	52.8
Class D 416 × 240	PartyScene	0.9	20.6	0.4	25.3	0.8	30.2	1.5	53.8
	RaceHorses	0.7	19.2	0.0	16.3	1.0	27.3	1.6	46.6
	BasketballPass	0.6	20.9	0.6	32.2	0.1	32.5	0.9	55.5
Class E 1280 × 720	BQSquare	0.5	18.2	0.6	27.8	0.4	33.9	1.2	57.3
	Blowingbubbles	0.5	14.4	0.3	18.8	0.5	22.1	1.1	42.6
	RaceHorses	0.3	11.1	0.4	22.3	0.6	17.7	1.1	35.9
Average	FourPeople	0.4	27.7	0.2	40.0	0.2	51.6	0.8	71.7
	Johnny	0.4	30.2	0.2	40.7	0.5	45.7	1.5	71.1
	KristenAndSara	0.4	34.7	0.2	42.5	0.6	51.6	1.4	75.6
		0.7	24.7	0.4	24.9	0.6	34.7	1.5	56.9

To evaluate the effectiveness of the proposed fast inter-frame encoding scheme, we compare our proposed method with four state-of-the-art fast encoding methods [13, 16, 20, 23] under the RA and LD conditions, respectively. We have tested our proposed method in different versions of HEVC reference software such as HM16.0, HM12.0, and HM10.0 and so on. We also test our method in different hardware platforms. We found that in those reference software and hardware platforms our proposed method have a same encoding performance. So the proposed method is independent of different versions of HEVC reference software and hardware platforms. Table 6 and 7 show the comparisons of encoding time reduction and BDBR loss for the proposed fast inter-frame encoding scheme and the other methods under the RA and LD conditions, respectively.

As shown in Tables 6 and 7, the performance of proposed overall fast inter-frame encoding scheme is better than those state-of-the-art fast encoding methods in [13, 16, 20, 23] in terms of encoding time reduction and BDBR loss for the RA and LD cases, respectively. In Table 6, under the RA condition the average encoding time reduction is 53.7% with 1.5% BDBR loss while Shen's method [13, 16] achieved time saving 47.7% with 0.7% BDBR loss and 40.7% with 1.6% BDBR loss, respectively. For test sequence "Traffic" the encoding time reduction can reach up to 63.2% with 2.0% BDBR loss and the best performance in [13] is 52.2% time reduction with 1.4% BDBR loss and in [16] is 59.3% with 0.9% BDBR loss. As we can see the proposed overall fast inter-frame encoding scheme outperforms the other two methods in [13, 16] in encoding time saving. Although the BDBR loss is higher than the method in [16] it also small enough that can be accept. In Table 7, we compare our proposed scheme with the method in [20, 23] under the LD condition. The average encoding time reduction is 54.9% and

**Table 6** Result comparison of proposed overall method with recent works for RA condition

Class resolution	Sequences	Proposed overall fast encoding scheme		Fast encoding scheme in Shen et al. [16]		Fast encoding method in Shen et al. [13]	
		BDBR(%)	$\Delta T$ (%)	BDBR(%)	$\Delta T$ (%)	BDBR(%)	$\Delta T$ (%)
Class A 2560 × 1600	Traffic	2.0	63.2	0.9	59.3	1.5	52.1
	PeopleOnStreet	1.5	48.4	1.1	38.3	1.4	30.4
Class B 1920 × 1080	Kimono1	2.5	58.6	0.7	42.8	0.8	45.3
	ParkScene	1.7	61.8	1.1	56.4	1.4	47.2
	Cactus	1.6	60.3	0.9	54.7	1.2	46.6
	BasketballDrive	2.1	57.5	0.6	48.7	1.4	46.0
	BQTerrace	1.5	60.0	0.7	58.8	1.4	52.2
Class C 832 × 480	BasketballDrill	1.5	51.1	0.1	46.3	1.7	46.1
	BQMall	1.4	52.8	0.9	52.2	3.5	50.2
	PartyScene	1.5	53.8	0.8	42.7	1.2	42.5
	RaceHorses	1.6	46.6	1.0	34.6	1.9	27.9
Class D 416 × 240	BasketballPass	0.9	55.5	0.4	51.4	1.9	33.4
	BQSquare	1.2	57.3	0.5	53.5	1.4	37.2
	Blowingbubbles	1.1	42.6	0.6	44.0	0.9	31.5
	RaceHorses	1.1	35.9	0.9	31.5	0.8	21.2
	Average	1.5	53.7	0.7	47.7	1.6	40.7

**Table 7** The result comparison of proposed overall method with recent works for LD condition

Class resolution	Sequences	Proposed overall fast encoding scheme		Fast encoding scheme in Zhang et al. [23]		Fast encoding method in Xiong et al. [20]	
		BDBR(%)	$\Delta T$ (%)	BDBR(%)	$\Delta T$ (%)	BDBR(%)	$\Delta T$ (%)
Class B 1920 × 1080	Kimono1	2.8	58.7	0.9	45.9	2.4	36.7
	ParkScene	2.2	60.2	1.9	45.2	2.5	39.8
	Cactus	1.8	58.1	2.9	48.1	2.5	43.8
	BasketballDrive	1.7	59.5	2.5	47.8	4.6	46.0
	BQTerrace	2.0	58.3	2.2	54.4	1.6	31.9
Class C 832 × 480	BasketballDrill	1.4	52.6	2.0	46.8	1.7	44.2
	BQMall	1.6	50.6	2.7	50.2	1.9	42.4
	PartyScene	2.0	50.6	1.6	43.1	1.5	30.3
	RaceHorses	1.5	44.2	1.6	33.4	2.0	30.9
Class D 416 × 240	BasketballPass	1.2	49.1	2.7	34.9	1.2	44.5
	BQSquare	1.6	51.5	1.9	38.8	2.3	33.0
	Blowingbubbles	1.3	39.1	0.8	35.6	1.7	25.6
	RaceHorses	1.2	37.7	1.9	26.0	1.5	31.1
Class E 1280 × 720	FourPeople	1.5	68.6	2.2	64.2	2.3	53.6
	Johnny	2.5	68.4	1.9	59.5	2.1	50.8
	KristenAndSara	2.6	71.8	3.0	63.6	2.6	53.1
	Average	1.8	54.9	2.0	46.1	2.2	39.9

BDBR loss is 1.8% for our method which is better than 46.1% encoding time reduction with 2.0% BDBR loss in [23] and 39.9% encoding time reduction with 2.2% BDBR loss in [20]. From the experiment results we can see that our proposed fast inter-frame encoding scheme can achieve a better performance in encoding time reduction than these state-of-the-art methods and just bring a negligible BDBR loss.

## 5 Conclusion

In this paper, we propose a fast Inter-frame encoding scheme utilizes the mean edge gradient value to reflect the video contents complexity, the edge similarity and the edge movement to reflect the video contents similarity. According to the mean edge gradient value to fulfill the fast all  $2N \times 2N$  modes decision, and use edge similarity and edge movement to selectively use the spatiotemporal encoding parameters to reduce the encoder complexity of HEVC. There are three approaches, a fast all  $2N \times 2N$  modes decision use the mean edge gradient value, a fast CU depth level decision uses the spatiotemporal CUs depth levels with the edge similarity and the edge movement and a fast PU mode decision uses the spatiotemporal CUs RD costs with the edge movement. The experimental results show that the proposed fast method can reduce the computational costs which can be represented by encoding time reduction reach up to 53.7 and 54.9% on average only with average 1.5 and 1.8% BDBR losses for various test sequences under RA and LD conditions, respectively. Furthermore, it consistently outperforms those state-of-the-art fast encoding methods in [13, 16, 20, 23].



## References

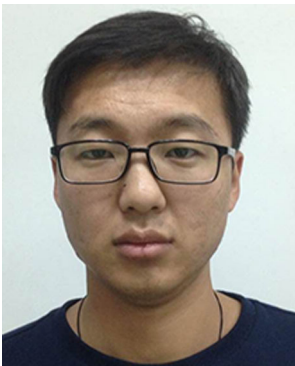
1. Ahn S, Lee B, Kim M (2015) A novel fast CU encoding scheme based on spatiotemporal encoding parameters for HEVC inter coding. *IEEE Trans Circ Syst Video Technol* 25(3):422–435
2. Bossen F (2013) Common HM test conditions and software reference configurations. Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Document: JCTVC-K1100, 12th Meeting, Geneva, CH
3. Bossen F, Bross B, Suhring K, Flynn D (2012) HEVC complexity and implementation analysis. *IEEE Trans Circ Syst Video Technol* 22(12):1685–1696
4. Cen Y, Wang W, Yao X (2015) A fast CU depth decision mechanism for HEVC. *Inf Process Lett* 115(9):719–724
5. Correa G, Assuncao PA, Agostini LV, da Silva Cruz LA (2015) Fast HEVC encoding decisions using data mining. *IEEE Trans Circ Syst Video Technol* 25(4):660–673
6. Fan X, Malone B, Yuan C (2014) Finding optimal Bayesian network structures with constraints learned from data. *Proc 30th Ann Conf Uncert Artif Intell (UAI-14)*: 200–209
7. Fan X, Yuan C, Malone BM (2014) Tightening bounds for Bayesian network structure learning. *AAAI*: 2439–2445
8. Kang J, Choi H, Kim J-G (2013) Fast transform unit decision for HEVC. *Imag Sign Process (CISP)*: 26–30
9. Lainema J, Bossen F, Han W-J, Min J, Ugur K (2012) Intra coding of the HEVC standard. *IEEE Trans Circ Syst Video Technol* 22(12):1792–1801
10. Lee A, Jun D, Kim J, Choi JS, Kim J (2014) Efficient inter prediction mode decision method for fast motion estimation in high efficiency video coding. *ETRI J* 36(4):528–536
11. Lee J, Kim S, Lim K, Lee S (2015) A fast CU size decision algorithm for HEVC. *IEEE Trans Circ Syst Video Technol* 25(3):411–421
12. Lee H, Shim HJ, Park Y, Jeon B (2015) Early skip mode decision for HEVC encoder with emphasis on coding quality. *IEEE Trans Broadcast* 61(3):388–397
13. Shen L, Liu Z, Zhang X, Zhao W, Zhang Z (2013) An effective CU size decision method for HEVC encoders. *IEEE Trans Multimed* 15(2):465–470
14. Shen X, Yu L, Chen J (2012) Fast coding unit size selection for HEVC based on Bayesian decision rule. *Picture Coding Sym (PCS)*: 453–456
15. Shen L, Zhang Z, Xi Z, An P, Liu Z (2015) Fast TU size decision algorithm for HEVC encoders using Bayesian theorem detection. *Signal Process Image Commun* 32:121–128
16. Shen L, Zhang Z, Liu Z (2014) Adaptive inter-mode decision for HEVC jointly utilizing inter-level and spatio-temporal correlations. *IEEE Trans Circ Syst Video Technol* 24(10):1709–1722
17. Sullivan GJ, Ohm J, Han W-J, Wiegand T (2012) Overview of the high efficiency video coding (HEVC) standard. *IEEE Trans Circ Syst Video Technol* 22(12):1649–1668
18. Xiong J, Li H, Meng F, Wu Q, Ngan KN (2015) Fast HEVC inter CU decision based on latent SAD estimation. *IEEE Trans Multimed* 17(12):2147–2159
19. Xiong J, Li H, Meng F, Zhu S, Wu Q, Zeng B (2014) MRF-based fast HEVC inter CU decision with the variance of absolute differences. *IEEE Trans Multimed* 16(8):2141–2153
20. Xiong J, Li H, Wu Q, Meng F (2014) A fast HEVC inter CU selection method based on pyramid motion divergence. *IEEE Trans Multimed* 16(2):559–564
21. 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
22. Zhang Y, Wang H, Li Z (2013) Fast coding unit depth decision algorithm for interframe coding in HEVC. *Proc DCC*: 53–62
23. Zhang Q, Zhao J, Huang X, Gan Y (2015) A fast and efficient coding unit size decision algorithm based on temporal and spatial correlation. *Optik* 126(21):2793–2798
24. 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):1709–1722
25. Zhong G, He X, Qing L, Li Y (2015) A fast inter-prediction algorithm for HEVC based on temporal and spatial correlation. *Multimed Tools Appl* 74:11023–11043
26. Zhou C, Zhou F, Chen Y (2013) Spatio-temporal correlation-based fast coding unit depth decision for high efficiency video coding. *J Electron Imag* 22(4):043001



**Zhiyao Yang** received the B.E. degree in Jilin University of Electronic Science and Engineering, Changchun, China, in 2012. Then he starts PhD in Jilin University. His research interests include Digital TV, Video Coding, 3D Video Processing, High Performance Computing.



**Shuxu Guo** received the B.Sc., M.Sc., and PhD degrees in electronic engineering from Jilin University, Changchun, China, in 1982, 1989, and 2006, respectively. Since 2000, is a professor in Electronic Science and Engineering at the Jilin University. His research interests are Signal Processing, Compressive Sensing, Communication Networks, Information Theory and so on. He has made some significant contributions to these research areas.



**Qinglong Shao** received the B.Sc. degree in Changchun University of Science and Technology. Since 2013, is a graduate student in Electronic Science and Engineering at the Jilin University. His research interests are Video Coding, High Performance Computing.