International Journal of Innovative Computing, Information and Control
Volume **14**, Number **1**, February **2018**

ICIC International ©2018 ISSN 1349-4198
pp. **309–322**

AN EFFICIENT COMPLEXITY REDUCTION ALGORITHM FOR CU SIZE DECISION IN HEVC

Xiantao Jiang[1], Xiaofeng Wang[1], Tian Song[2], Wen Shi[2]
Takafumi Katayama[2], Takashi Shimamoto[2] and Jenq-Shiou Leu[3]

[1]Department of Information Engineering
Shanghai Maritime University
No. 1550, Harbour Rd., Haigang Ave., Pudong New Dist., Shanghai 201306, P. R. China
xtjiang@shmtu.edu.cn

[2]Department of Electrical and Electronics Engineering
Tokushima University
2-24, Shinkura-cho, Tokushima 770-8501, Japan
tiansong@ee.tokushima-u.ac.jp

[3]Department of Electrical and Computer Engineering
National Taiwan University of Science and Technology
No. 43, Keelung Rd., Sec. 4, Da'an Dist., Taipei City 10607, Taiwan
jsleu@mail.ntust.edu.tw

Received May 2017; revised September 2017

Abstract. *In brief, this work focuses on reducing the encoding complexity with less than 1% encoding efficiency loss for low delay (LD) and random access (RA) profiles in HEVC. An efficient CU (coding unit) size decision algorithm based on probabilistic graphical models is proposed for HEVC inter coding, which contains two methods: the CU size early termination (CUET) decision method and the CU size early skip (CUES) decision method. The CU pruning is modeled as a binary classification problem based on the Naive Bayes (NB) model. Furthermore, a Markov random fields (MRF) model based method is presented to improve the algorithm performance. The difference from previous works is that residual flag in inter-coded CU and the neighboring information are used to determine the CU size. The offline learning method is used to obtain the statistical parameters. This presented approach can significantly reduce the encoding complexity. Furthermore, it can bring lower power cost for hardware implementation. The simulation experiment results demonstrate that this method can significantly reduce by 50.59% and 53.86% average encoding complexity under low delay and random access profiles, while the encoding efficiency can be reduced by 0.82% and 0.98% on average. Moreover, the rate-distortion (RD) performance of this method is nearly the same as HEVC reference software.*

Keywords: Video coding, CU size, Encoding efficiency, Encoding complexity

1. **Introduction.** High efficiency video coding (HEVC) is the latest video coding standard that is released in 2013 [1]. It is a hybrid coding model, and it achieves 50% bitrate saving compared with H.264/AVC. HEVC divides the picture into coding tree unit (CTU). Then a CTU is further divided into four coding units (CU) in a quad-tree structure. Prediction unit (PU) is a smaller unit, defined by partitioning the CU. Each inter coded PU has a set of motion parameters including motion vector (MV), picture index, and reference picture list. For each PU, there are three available modes: inter, skip and merge modes. In the partitioning unit, the partition modes are used to define the prediction unit for inter-coded CU. Partitioning modes include two square partition modes, two symmetric motion partition modes and four asymmetric motion partition modes. Usually, the

best motion vector is selected from certain motion vector prediction candidates which can minimize the rate-distortion cost.

In HEVC, the maximum CU size is 64×64, and each of this sub-CU can be divided into smaller sub-CUs recursively in the same quad-tree. As a result, the intra and inter predictions take the most encoding time in the whole encoding process. The implementation cost is unacceptable for both software and hardware implementation, especially for those real-time applications. Therefore, on the premise of guaranteeing the coding performance, reducing the encoding complexity is a key factor for the success of HEVC, and the need of large scale HD video application.

However, CU size is critical in HEVC, and more researches focus on CU fast decision algorithm. The purpose of this kind of method is to reduce the encoding complexity of CU size decision, on the premise of guaranteeing the coding quality. There are differences between intra prediction and inter prediction. The intra CU size fast decision algorithm mainly targets the CU size decision of intra mode prediction, while inter CU size fast decision algorithm mainly targets the CU size decision of P frame and B frame. Furthermore, the intra CU size fast decision algorithm is to determine the CU size and intra prediction mode beforehand by evaluating the CU texture complexity or on the basis of the CU depth information. The inter CU size fast decision algorithm is to determine the CU size and inter PU mode beforehand on the basis of the neighboring CU depth or the middle encoding parameters. In this work, the proposed algorithm mainly focuses on the inter CU size fast decision.

Many previous works have major attention to reduce encoding complexity of HEVC encoder [2-23]. According to different usage of information, there are three categories for the CU size fast decision: (1) based on the middle encoding parameters, (2) based on the neighboring CU depth, and (3) based on the rate-distortion cost (RD-cost). The detailed description of these methods is as follows.

The main idea based on the middle encoding parameters of inter prediction fast selection algorithm is that the middle encoding parameters (such as motion vector (MV), coded block flag (CBF), MV difference (MVD), TU size, and SAO) can effectively reflect the CU motion compensation effect of inter prediction. These methods can effectively determine the current CU coding mode by using these encoding parameters. The representative works have the following. Ahn et al. propose the inter CU fast decision algorithms based on the spatial and temporal encoding information [2,3]. These methods use SAO parameters to estimate the CU texture complexity, and use MV, TU size and CBF to estimate the CU motion complexity. Shen et al. propose a fast CU size decision approach based on Bayes rule [4], and the feature parameters include sum of absolute transformed difference (SATD) and MV that are used to decide the CU coding mode based on Bayes rule. Kim et al. propose an early inter mode termination algorithm [5], and this method mainly uses the MVD and CBF information to check the skip mode. Ahn et al. use the zero block detection and MVD information to early terminate the inter prediction CU split process [6]. Pan et al. propose a merge mode detection of inter CU size decision [7], and this approach uses the zero block detection and inter motion estimation information to early terminate the merge mode. The above mentioned methods mainly use the middle encoding parameters, and selecting the optimal mode depending on these parameters reflects the effect of CU coding. However, these methods lack precision for CU size decision.

The main idea based on the neighboring CU depth of inter CU fast selection algorithm is that it uses the neighboring CU depth information to deduce the current CU depth range. Because the neighboring CU motion complexity and encoding size are close to the current CU motion complexity and size, which can decide the current CU coding mode faster

by using the neighboring CU encoding information. The representative works have the following. Shen et al. propose a fast inter CU depth range selection algorithm [8], and this approach uses the spatial neighboring CU depth information and temporal neighboring CU depth information to deduce the current CU depth range. Zhang et al. use the depth similarity of current CTU and neighboring spatial-temporal CTU to zoom out the depth range [9]. Leng et al. propose a fast inter CU depth determination method [10], and this method uses the neighboring spatial-temporal CU depth information in frame level and CU level to skip some CU rate-distortion optimization process. Lee et al. propose a fast mode selection algorithm based on neighboring spatial-temporal CU information [11]. Similarly, Correa et al. use spatial-temporal CU depth to estimate the current CU depth [12]. Mu et al. propose a fast CU depth decision scheme to reduce the encoder complexity for HEVC [13]. The method is used to predict the CU depth based on the support vector machine (SVM) model. The robustness of these methods is not high, because it depends on the consistency of the texture and motion characteristics between neighboring CU and the current CU.

The main idea based on the RD-cost of inter CU fast selection algorithm is that the RD-cost can reflect the final motion compensation effect of current CU significantly, and it is able to determine early coding mode by calculating part of the pattern of the RD-cost. The representative works have the following. Tan et al. propose a fast inter CU selection algorithm [14]. Lee et al. propose a fast inter CU mode selection method based on the RD-cost among encoding process [15], and the RD-cost of skip mode and the optimal RD-cost after executing skip or merge and $2N \times 2N$ mode are used to decide the CU encoding mode. Cassa et al. propose an approach to determine whether it needs to skip or terminate the current encoding process by comparing the RD-cost with the setting threshold [16]. Vanne et al. propose a fast inter mode selection algorithm [17]. This method analyzes the distribution of inter prediction mode and the relationship of different prediction modes, and judgment condition whether PU mode needs to execute or not is proposed. Zhang et al. propose an RD-cost based fast CU depth decision algorithm [18], and the method includes a three-output joint classifier to control the risk of false prediction. The optimal mode of above methods depends on the threshold of RD-cost. However, this threshold is relative to the image texture and motion intensity, and the threshold is dynamic variation. Thus, the implementation cost is high, because it costs much time to calculate the RD-cost.

The above three methods have advantages and disadvantages. The advantages of the middle encoding parameters method are that the feature extraction is straightforward and it does not need the extra computational complexity, while the accuracy of this method is low. The advantages of the neighboring CU depth method are that this method is simple and is easy to achieve; however, it depends on the consistency of the neighboring CU. The advantages of the RD-cost method are that the RD-cost threshold determination method is easy to implement, while the cost of RD-cost calculation is very time-consuming and the robustness is not high. Although Jiang et al. propose a fast CU size decision algorithm based on the RD-cost and the middle encoding parameters to reduce the encoding complexity in [22,23], they have some limitations. (1) This method cannot achieve the trade-off between the encoding efficiency and encoding complexity, and the average loss of encoding efficiency is greater than 1%. (2) The more effective information of neighboring CU has not been used to determine the CU size.

As a summary, the state-of-the-art algorithms cannot achieve a better trade-off between encoding complexity and encoding efficiency. For high-definition applications corresponding to LD profile, the loss encoding efficiency should be negligible. Meanwhile, for network applications corresponding to RA profile, the encoding efficiency can be sacrificed a little

bit more to reduce the encoding complexity. Thus, this work focuses on reducing the encoding complexity with less than 1% encoding efficiency loss for low delay (LD) profile and random access (RA) profile in HEVC. An efficient CU size decision algorithm based on probabilistic graphical models (PGM) is proposed.

The remainder of this paper is organized as follows. In Section 2, we review the rate distortion model and investigate the probability distribution of CU splitting or non-splitting. In Section 3, the CU size decision based on probabilistic graphical models is introduced to reduce the encoding complexity. In Section 4, we present some experimental results. Finally, Section 5 concludes this paper.

2. **Observation and Statistical Analysis.** In HEVC reference software HM, a two-step rate distortion optimization (RDO) method is used for mode decision. At first, to save computation overhand, a fast RDO is used for early termination, and fast RDO selects the motion vectors and modes for inter prediction. The minimized low-complexity RD-cost function J_{pred} is defined as

$$\min J_{pred} = D_{pred} + \lambda_{pred} \times R_{pred} \tag{1}$$

where D_{pred} represents the distortion between the original block and reference block, R_{pred} represents the number of coding bits, and λ_{pred} is the Lagrange multiplier. Then, a full RDO is used for the final decision. The full RDO is determinated by CABAC bit rate and distortion cost, and the minimized full RD-cost function J_{mode} is defined as

$$\min J_{mode} = (SSE_{luma} + \omega_{chroma} \times SSE_{chroma}) + \lambda_{mode} \times R_{mode} \tag{2}$$

where SSE_{luma} and SSE_{chroma} are on behalf of the sum of square error (SSE) between the original and reconstructed luma and chroma blocks, R_{mode} represents the number of the coding bits, and ω_{chroma} is weighting factor. However, the cost of full RDO is high in real-time HEVC encoder. Similarly, λ_{mode} is Lagrange multiplier, and the definition λ_{mode} and the relationship between λ_{mode} and λ_{pred} are that:

$$\lambda_{mode} = \alpha \times W_k \times 2^{\frac{(QP-12)}{3.0}} \tag{3}$$

$$\lambda_{pred} = \sqrt{\lambda_{mode}} \tag{4}$$

where QP is quantization parameter, and W_k is weighting factor.

$$\alpha = \begin{cases} 1.0 - Clip3(0.0, 0.5, 0.05 \times \text{number\_of\_B\_frames}); & \text{for referenced pictures} \\ 1.0; & \text{for non-referenced pictures} \end{cases}$$

where

$$Clip3(x, y, z) = \begin{cases} x; & z < x \\ y; & z > y \\ z; & \text{otherwise} \end{cases}$$

The CBF is an important factor for deciding the CU size decision [27]. When CBF equals zero, the image texture tends to be more smooth. Therefore, the current CU tends not to be split. On the contrary, when CBF equals one, the image texture tends to be complex. Therefore, the current CU tends to be split.

In HEVC, based on RDO model, the RD-cost is calculated to decide whether the current CU splits or not. Therefore, the current CU non-splitting or splitting is formulated as a binary classification problem. This two-class event is denoted by discrete random variable y, which is defined as

$$y = \begin{cases} 0 & \text{if CU non-splitting} \\ 1 & \text{if CU splitting} \end{cases}$$

Extensive experiments based on HEVC reference software (HM12.0) are to investigate the probability distribution of CU splitting or non-splitting with different class sequences, and the results are shown in Table 1. The sequences have different resolutions: Traffic (2560×1600), BQTerrace (1920×1080), Vidyo1 (1280×720), BQMall, and BlowingBubbles (416×240). The profile is LD (lowdelay), and CU size is 64×64 when QP (quantization parameter) is 32. The result is shown that the probability of CU non-splitting is high for high resolution. On the contrary, the probability of CU splitting is high for low resolution. Similar to the same simulation environment, Table 2 represents the conditional probability distribution of $p(x_{cbf}|y)$ with different resolution sequences, where x_{cbf} represents the value of CBF, and $p0 = p(x_{cbf} = 0|y = 0)$, $p1 = p(x_{cbf} = 1|y = 1)$. It can be seen that the conditional probability of $x_{cbf} = 0$ is high for CU non-splitting. On the contrary, the conditional probability of $x_{cbf} = 1$ is high for CU splitting.

TABLE 1. Probability distribution of the CU non-splitting (NS) and splitting (S)

Sequence Resolution	**Sequence**	**NS**	**S**
2560×1600	**Traffic**	0.60	0.40
1920×1080	**BQTerrace**	0.70	0.30
1280×720	**Vidyo1**	0.82	0.18
High Resolution	**Average**	0.71	0.29
832×480	**BQMall**	0.16	0.84
416×240	**BlowingBubbles**	0.48	0.52
Low Resolution	**Average**	0.32	0.68

TABLE 2. Conditional probability distribution of $p(x_{cbf}|y)$

Sequence Resolution	**Sequence**	$p0$	$p1$
2560×1600	**Traffic**	0.96	0.73
1920×1080	**BQTerrace**	0.97	0.54
1280×720	**Vidyo1**	0.96	0.61
High Resolution	**Average**	0.96	0.63
832×480	**BQMall**	0.91	0.88
416×240	**BlowingBubbles**	0.88	0.70
Low Resolution	**Average**	0.90	0.79

In addition, based on some observations from experiments with many sequences, the RD-cost probability density function (*pdf*) of the CU non-splitting and CU splitting obeys Gaussian distribution [15].

3. **Proposed CU Size Decision for Inter Prediction.** In this section, an effective CU size decision algorithm is presented in the HEVC inter prediction. However, this method is different from Shen et al.'s [4] and Lee et al.'s [15] work. Firstly, the CBF flag is used to make CU size decision. Secondly, the decision strategy is based on Naive Bayes model. Furthermore, this statistical parameters are estimated by using offline learning and online learning methods.

3.1. **Naive Bayes based CU size decision algorithm.** Thus, for the random variable y of the current CU splitting or non-splitting, the probability density function (*pdf*) $p(y)$ follows a discrete Bernoulli (p) distribution, which is defined as

$$p(y) = \begin{cases} p & \text{if } y = 0 \\ 1 - p & \text{if } y = 1 \end{cases}$$

where p is the probability of CU non-splitting.

For the binary classification problem $y = \{0, 1\}$, the assumption is that the features of CU are comprised of $x = \{x_1, x_2, \ldots\}$, and these features are independent of each other. Having fit these parameters, to make a prediction on CU non-splitting or splitting with features $\{x_1, x_2, \ldots\}$, the CU termination decision rule is

$$\begin{cases} p(y = 0|x) > p(y = 1|x) & \text{CU non-splitting is made} \\ \text{else} & \text{CU splitting is made} \end{cases}$$

where the class-conditional probability density function $p(y|x)$ is calculated based on Naive Bayes (NB) model

$$p(y|x_1, x_2, \ldots) = \frac{p(x_1, x_2, \ldots|y)p(y)}{p(x_1, x_2, \ldots)} = \frac{\prod_{i=1}^{n} p(x_i|y)p(y)}{p(x_1, x_2, \ldots)} \tag{5}$$

where the features x_i are the coded block flag (CBF) and RD cost of partition $2N \times 2N$, denoted as x_1 and x_2. The prior probability function $p(x_1|y)$ is modeled using a discrete Bernoulli (ϕ) distribution, which are defined as

$$p(x_1|y) = \begin{cases} \phi & \text{if } y = 0 \\ 1 - \phi & \text{if } y = 1 \end{cases}$$

The prior probability function $p(x_2|y)$ is modeled using the Gaussian distribution. The model is

$$x_2|y = 0 \sim N\left(\mu_0, \sigma_0^2\right)$$
$$x_2|y = 1 \sim N\left(\mu_1, \sigma_1^2\right)$$

where the parameters (μ_0, σ_0), (μ_1, σ_1) are mean vectors and covariance matrix of CU non-splitting and splitting, respectively. It is noted that these parameters are estimated by the maximum likelihood estimation. Thus, the prior probability of $p(x_2|y)$ are defined as

$$p(x_2|y = 0) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(x_2-\mu_0)^2}{2\sigma_0^2}} \tag{6}$$

$$p(x_2|y = 1) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x_2-\mu_1)^2}{2\sigma_1^2}} \tag{7}$$

Actually, the evidence probability $p(x)$ is the constant. In order to make a prediction, we update the prior distribution to the posterior

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)} \propto p(x|y)p(y) = \left(\prod_{i=1}^{n} p(x_i|y)\right) p(y) \tag{8}$$

Thus, the decision function $D(y)$ is defined as

$$D(y) = \left(\prod_{i=1}^{n} p(x_i|y)\right) p(y) \tag{9}$$

After the introduction of classifier, the CU splitting and non-splitting are decided by Naive Bayes classifier. Therefore, the approach includes two methods: CU early termination (CUET) decision and CU early skip (CUES) decision. CUET decision is to terminate the CU further splitting, and CUES decision is to skip the CU mode selection in the current depth and go to the next depth of the CU.

Thus, when the context of image tends to be smooth, the probability of CU non-splitting is higher than the probability of CU splitting. The condition of CUET decision is that

$$CUET\ condition \begin{cases} CBF = 0 \\ D(y=0) > D(y=1) \end{cases}$$

On the other hand, when the context of image tends to be complex, the probability of CU splitting is higher than the probability of CU non-splitting. The condition of CUES decision is that:

$$CUES\ condition \begin{cases} CBF = 1 \\ D(y=1) > D(y=0) \end{cases}$$

3.1.1. *Statistical parameter estimation.* Table 3 summarizes the statistical parameters. In our approach, the statistical parameters (mean, standard deviation and prior) are estimated by the offline learning method [26].

TABLE 3. The lookup table of estimation parameter (learning parameter: LP, derived parameter: DP)

LP	DP	Description
p	$p(y)$	Probability of CU non-splitting
(μ_i, σ_i)	$p(x_{cost}\|y)$	Conditional probability of RD cost
ϕ	$p(x_{cbf}\|y)$	Conditional probability of CBF

The statistical parameters are estimated by using a non-parametric estimation with offline learning [26], and are stored in a lookup table (LUT0). It is noticed that the statistical parameters are varied as the CU depth, QP and resolution changed. Thus, we select the sequences: Traffic (2560×1600), BQTerrace (1920×1080), Vidyo4 (1280×720), BQMall (832×480), and BlowingBubbles (416×240) for non-parametric estimation. In the inter prediction stage by using the presented algorithm, the statistical parameters are indexed by CU depth, QP and resolution.

3.1.2. *Overall algorithm.* By joining the CU termination and skip algorithm, the flowchart of presented overall algorithm based on offline learning or online learning is shown as Figure 1. This overall algorithm can be divided into five steps.

Step 1: Start CU size decision with motion estimation process.

Step 2: Look up the statistical parameters in LUT0.

Step 3: When the current mode is $2N \times 2N$ and CBF is zero, calculate $D(y=0)$ and $D(y=1)$ as Formula (9). If $D(y=0) > D(y=1)$, it allows terminating all the remaining RD cost computing.

Step 4: When the current mode is $2N \times 2N$ and CBF is 1, if $D(y=1) > D(y=0)$, it skips the RD-cost computing for remaining CUs in the same depth and goes to the next depth of the CU.

Step 5: If the current depth is less than the maximal depth, depth = depth + 1 and repeat Steps 2 and 3. Otherwise, the best CU size is determined.

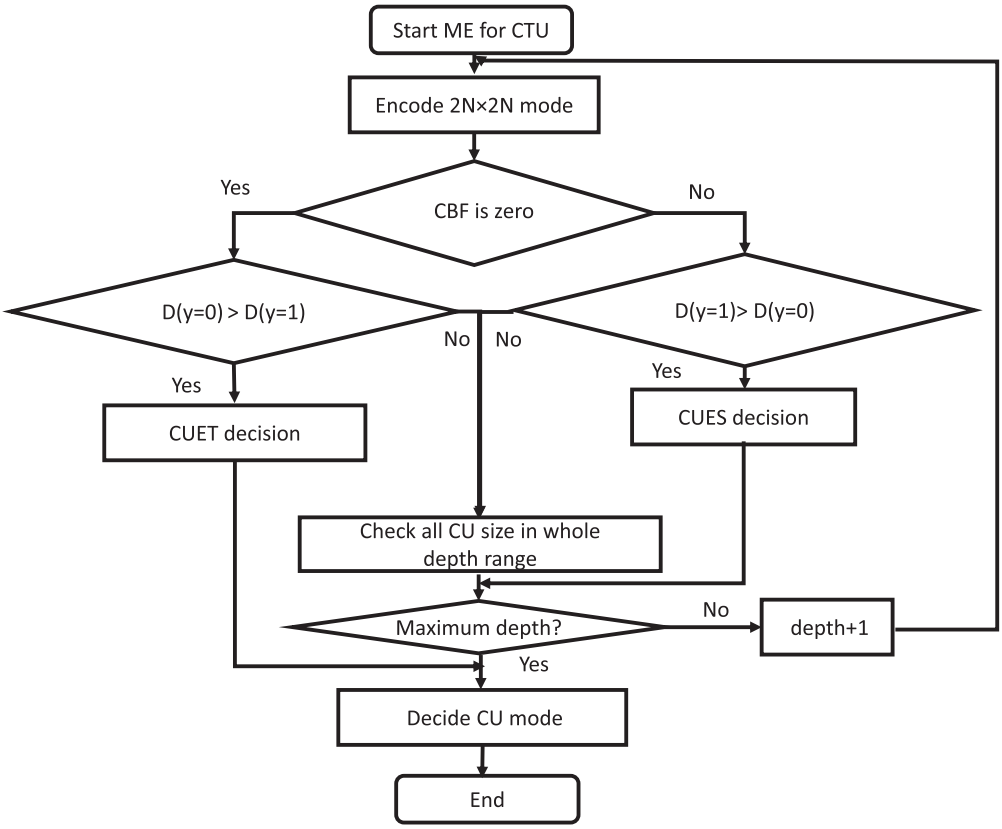


FIGURE 1. Flowchart of the proposed overall algorithm

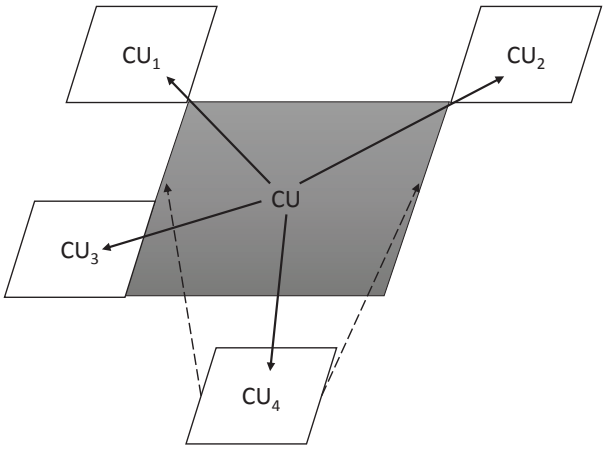


FIGURE 2. The neighborhood system of the current CU

3.2. **Markov random fields based CU size decision algorithm.** In HEVC, there is correlation between current CU and neighborhood CU. In order to utilize the spatial-temporal correlation, the four neighborhood system M is defined as

$$M = \{CU_1, CU_2, CU_3, CU_4\}$$

As Figure 2 shown, CU_1, CU_2, CU_3 denote the spatially adjacent CUs of the current CU, and CU_4 denotes the temporally adjacent CU of the current CU.

Whereas, the prior distribution $p(y)$ can be confirmed by the probabilistic graphical model: Markov random fields (MRF) [8].

$$p(y) = \frac{1}{Z} \exp\left(-\sum_{j \in M} V_j(X_j)\right) \tag{10}$$

From the physicists, this is the Gibbs distribution with interaction potential $\{V_j, j \in M\}$, energy $U = \sum_j V_j$, and partition function of parameters Z. Configurations of lower energies are more likely, whereas high energies correspond to low probabilities. The CU size decision is a binary classification problem, and the binary problem can be modeled by a simple ISING-MRF model [28]

$$V_j(X_j) = -\beta \times (X_0 \times X_j) \tag{11}$$

where the X_0 denotes the flag of the CU_0 splitting or non-splitting, and X_j denotes the flag of the CU_j splitting or non-splitting in neighborhood system M. β is coupling factor, which indicates the strength of CU correlation with neighborhood system M. Thus, the prior $p(y)$ exhibits a factorized form

$$p(y) \propto \exp\left(-\sum_{j \in M} -\beta \times (X_0 \times X_j)\right) \tag{12}$$

Take log function of the posterior $p(y|x)$, and it can be written as

$$\ln p(y|x) \propto \left[C_1 - \frac{1}{2\sigma_i^2}(x_2 - \mu_i)^2 - \sum_{j \in M} -\beta \times (X_0 \times X_j)\right] \tag{13}$$

where constant $C_1 = \ln p(x_1|y)$, $i = 0, 1$. Then when CU_0 neighborhood system M is valid, the decision function $D(y)$ can be defined as

$$D(y) = \left[C_1 - \frac{1}{2\sigma_i^2}(x_2 - \mu_i)^2 - \sum_{j \in M} -\beta \times (X_0 \times X_j)\right] \tag{14}$$

However, when CU_0 neighborhood system M is invalid, the $p(y)$ can be confirmed by Bernoulli (ψ) distribution. The decision function can be rewritten as

$$D(y) = \left[C_1 - \frac{1}{2\sigma_i^2}(x_2 - \mu_i)^2 + \ln p(y)\right] \tag{15}$$

Through the above analysis, the improved CU size decision algorithm based on ISING-MRF model includes CUET decision and CUES decision. The condition of CUET decision is: CBF = 0, and $D(y = 0) > D(y = 1)$. The condition of CUES decision is: CBF = 1, and $D(y = 1) > D(y = 0)$.

4. **Experimental Results.** The proposed algorithm is implemented and verified based on HEVC test model HM12.0. The test conditions are set to evaluate the performance of the presented algorithm at different profiles. The quantization parameters (QP_i) are set to 22, 27, 32 and 37, respectively. The coupling factor β ($0 < \beta < 1$) indicates the strength of CU correlation with neighborhood system M, and β is set to 0.5 and 0.75 in this work.

The performance of this algorithm is evaluated Bjontegarrd Delta bitrate (BR) [29], peak-signal-to-noise ratio (PSNR). The average time saving (TS) is defined as

$$TS(\%) = \frac{1}{4} \sum_{i=1}^{i=4} \frac{T_{HM}(QP_i) - T_{pro}(QP_i)}{T_{HM}(QP_i)} \times 100\% \tag{16}$$

where $T_{HM}(QP_i)$ and $T_{pro}(QP_i)$ are the encoding time by using the HEVC reference software and the presented method with different QP_i.

4.1. **Performance of Naive Bayes based CU size decision algorithm.** Table 4 shows the results of the Naive Bayes based method compared to HEVC reference software. The third and fourth columns in the table show the performance under the low delay (LD) profile. From the experimental results, it can be seen that, in the aspect of encoding complexity, the method can save 50.24% encoding time, while the encoding efficiency can be reduced by 1.36%. Thus, in the low delay profile, this method almost does not affect the encoding quality while the encoding complexity can be reduced significantly.

TABLE 4. The performance of the Naive Bayes based CU size decision algorithm

		LD		RA	
Class	**Sequence**	**BR**	**TS**	**BR**	**TS**
2560 × 1600	**Traffic**	1.23	55.29	1.63	56.17
	SteamLocomotive	0.40	51.56	0.74	55.86
1920 × 1080	**Kimono**	1.90	41.58	2.35	44.82
	ParkScene	1.15	52.5	1.33	55.77
	Cactus	1.56	47.66	1.78	49.56
	BasketballDrive	4.01	44.66	4.94	47.35
	BQTerrace	0.72	54.08	1.01	54.31
1280 × 720	**Vidyo1**	1.65	64.54	1.70	63.80
	Vidyo3	1.30	61.83	1.41	61.95
	Vidyo4	1.14	65.11	1.46	62.74
High Res.	**Average**	**1.51**	**53.81**	**1.83**	**55.13**
832 × 480	**BasketballDrill**	2.14	45.59	1.97	49.20
	BQMall	1.00	49.20	1.16	56.45
	PartyScene	0.67	39.53	0.96	50.25
	RaceHorses	1.27	36.78	1.79	43.62
416 × 240	**BasketballPass**	1.06	53.02	1.50	54.51
	BQSquare	0.34	47.58	0.65	53.40
	BlowingBubbles	1.50	43.56	1.45	49.32
Low Res.	**Average**	**1.14**	**45.04**	**1.35**	**50.96**
Average		1.36	50.24	1.64	53.48

The fifth and sixth columns in the table show the performance under the random access (RA) profile. In the aspect of encoding complexity, the method can save 53.48% encoding time, while the encoding efficiency can be reduced by 1.64%. Thus, in the low delay profile, this method almost does not affect the encoding quality while the encoding complexity can be reduced significantly.

The results demonstrate that this method can significantly reduce the encoding time, and the time saving in high resolution is higher than in the low resolution applications. Furthermore, for 1280 × 720 resolution, the time saving is more than 60% in the LD and RA profiles.

4.2. **Performance of Markov random fields based CU size decision algorithm.** The performance of the improvement CU size decision method is shown as Table 5. The third to sixth columns in the table show the performance of the improvement method, when β is set to 0.5. The third and fourth columns in the table show the performance under LD profile. From the experimental results, it can be seen that, in the aspect of

TABLE 5. The performance of the Markov random fields based CU size decision algorithm

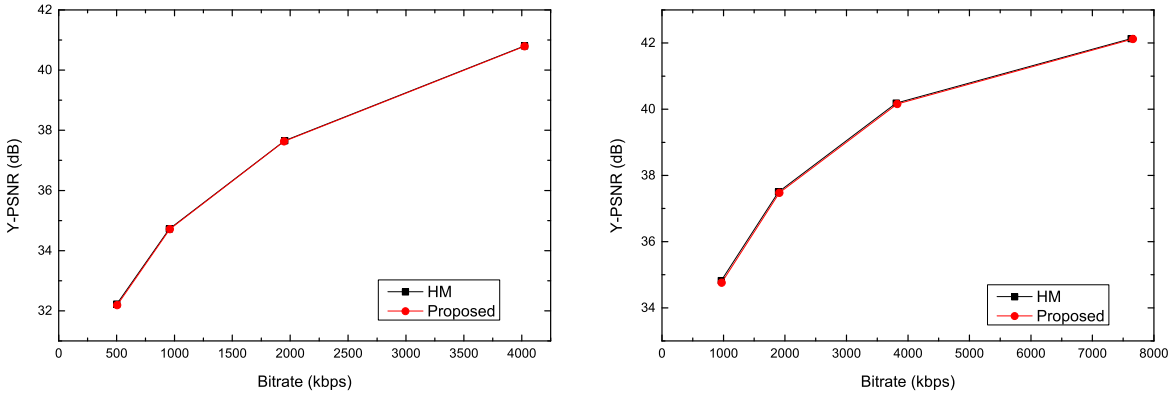
		$\beta = 0.5$				$\beta = 0.75$			
		LD		**RA**		**LD**		**RA**	
Class	**Sequence**	**BR**	**TS**	**BR**	**TS**	**BR**	**TS**	**BR**	**TS**
2560 × 1600	**Traffic**	0.85	54.72	1.07	59.99	0.89	55.29	1.06	56.17
	SteamLocomotive	0.22	50.8	0.78	58.39	0.22	51.56	0.78	55.86
1920 × 1080	**Kimono**	1.57	41.06	1.83	49.02	1.49	41.58	1.84	44.82
	ParkScene	0.99	52.19	1.15	58.54	1.07	52.5	1.10	55.77
	Cactus	1.05	46.84	1.36	53.68	0.98	47.66	1.36	49.56
	BQTerrace	0.73	53.76	0.69	57.8	0.71	54.08	0.70	54.31
1280 × 720	**Vidyo1**	0.95	64.47	1.87	66.63	0.98	64.54	0.86	63.80
	Vidyo3	0.72	61.88	0.63	64.69	0.86	61.83	0.59	61.95
	Vidyo4	0.61	64.77	1.01	65.26	0.32	65.11	1.09	62.74
High Res.	**Average**	**0.85**	**54.50**	**1.06**	**59.25**	**0.83**	**54.91**	**1.04**	**56.10**
832 × 480	**BasketballDrill**	0.59	46.47	0.72	49.92	0.54	45.59	0.74	49.2
	BQMall	0.94	50.31	1.06	55.56	0.74	49.20	0.98	56.45
	PartyScene	0.66	41.09	0.78	48.63	0.59	39.53	0.75	50.25
	RaceHorses	0.88	37.93	1.40	42.79	0.94	36.78	1.39	43.62
416 × 240	**BasketballPass**	1.02	52.83	0.85	53.17	0.81	53.02	0.66	54.51
	BQSquare	0.53	47.77	0.56	52.04	0.58	47.58	0.55	53.40
	BlowingBubbles	1.47	43.16	1.27	47.90	1.33	43.56	1.18	49.32
Low Res.	**Average**	**0.87**	**45.65**	**0.95**	**50.00**	**0.79**	**45.04**	**0.89**	**50.96**
Average		0.86	50.63	1.01	55.25	0.82	50.59	0.98	53.86

encoding complexity, the method can save 50.63% encoding time, while the encoding efficiency can be reduced by 0.86%. The fifth and sixth columns in the table show the performance under the RA profile. In the aspect of encoding complexity, the method can save 55.25% encoding time, while the encoding efficiency can be reduced by 1.01%. Thus, in the low delay profile, this method almost does not affect the encoding quality while the encoding complexity can be reduced significantly.

The seventh to tenth columns in the table show the performance of the improvement method, when β is set to 0.75. The seventh and eighth columns in the table show the performance under the LD profile. From the experimental results, it can be seen that, in the aspect of encoding complexity, the method can save 50.59% encoding time, while the encoding efficiency can be reduced by 0.82%. The ninth and tenth columns in the table show the performance under the RA profile. In the aspect of encoding complexity, the method can save 53.86% encoding time, while the encoding efficiency can be reduced by 0.98%. Thus, in the low delay profile, this presented method almost does not affect the encoding quality while the encoding complexity can be reduced significantly.

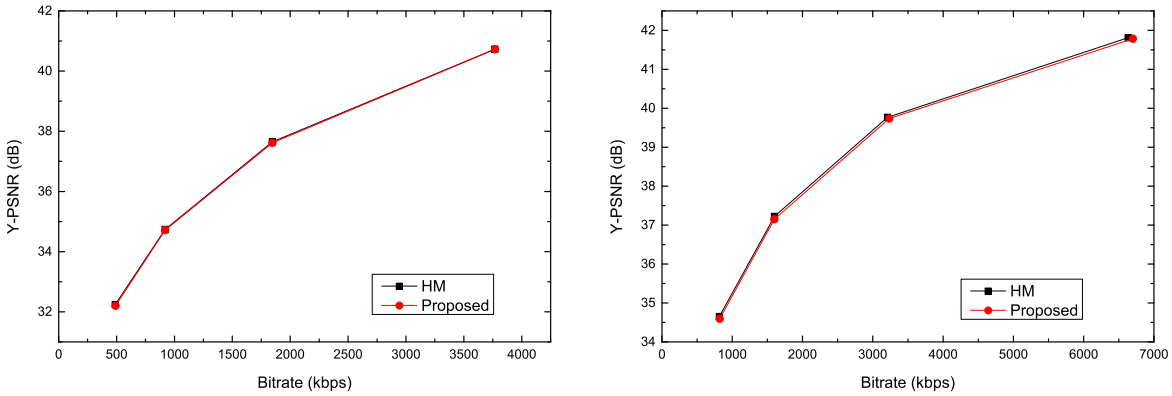
Through the analysis of experimental results, it is noted that the coupling factor β is a sensitivity parameter. When the value of β changes from 0.5 to 0.75, the BR and TS are decreased. In order to achieve the target that BR is less than 1%, a suitable value of β is set to 0.75 in this work.

Comparing Table 4 and Table 5, the encoding complexity of the improvement method is almost the same as the encoding complexity of the Naive Bayes based method, while the encoding efficiency of the improvement is better than the encoding efficiency of the Naive Bayes based method. That is, the improvement method can make a better trade-off between the encoding complexity and encoding efficiency.



(a) BasketballDrill (b) Kimono

FIGURE 3. R-D curve of the improvement method ($\beta = 0.75$, low delay)



(a) BasketballDrill (b) Kimono

FIGURE 4. R-D curve of the improvement method ($\beta = 0.75$, random access)

To evaluate the steady performance, the R-D curves of the typical sequences are as shown in Figure 3 and Figure 4 for the LD and RA profiles. It can be noticed that, no matter in high bitrate or in low bitrate, the R-D performance of the proposed method is almost similar to the HEVC reference software.

4.3. **Comparison with previous work.** Furthermore, the performance of the MRF based method is compared to the previous work [4,8,15,18,24,25] with $\beta = 0.75$, and the results are shown in Table 6. Shen et al.'s method [4] is based on the middle encoding parameters of inter prediction. Shen et al.'s [8] is based on the neighboring CU depth of inter CU fast selection. Zhang et al.'s method [18] and Lee et al.'s method [15] are based on the RD-cost of inter CU fast selection. Xiong et al.'s methods [24,25] are based on the RD-cost and the middle encoding parameters of inter prediction.

It can be seen from the comparison of experimental results that, no matter in high resolution or low resolution, the encoding complexity of the MRF based method can be reduced, significantly; meanwhile, the encoding efficiency of this method is better than previous work. Moreover, our approach can reduce the encoding complexity with less than 1% for LD and RA profiles when β is set to 0.75, respectively.

TABLE 6. The performance of the proposed method compared with previous work

	Method	(BR, TS)		
		High Res.	**Low Res.**	**Average**
LD	**Proposed**	(0.83, 54.91)	(0.79, 45.04)	(0.82, 50.59)
	Shen et al.'s [8]	(0.97, 47.11)	(1.33, 33.89)	(1.15, 41)
	Lee et al.'s [15]	(1.31, 69)	(1.13, 53)	(1.22, 61)
	Zhang et al.'s [18]	(2.41, 62.59)	(1.55, 40.31)	(1.98, 51.45)
	Xiong et al.'s [24]	(2.78, 66.13)	(1.6, 53.21)	(2.19, 59.67)
	Xiong et al.'s [25]	(2.59, 44.40)	(1.83, 36.26)	(2.21, 40.33)
RA	**Proposed**	(1.04, 56.10)	(0.89, 50.96)	(0.98, 53.86)
	Shen et al.'s [4]	(1.25, 51.05)	(1.33, 38.61)	(1.35, 44.7)
	Shen et al.'s [8]	(1.30, 45.25)	(1.65, 38.78)	(1.49, 42)
	Lee et al.'s [15]	(1.49, 65.43)	(1.37, 58.57)	(1.43, 62)
	Xiong et al.'s [24]	(3.3, 69.24)	(2.18, 57.10)	(2.74, 63.17)

5. **Conclusion.** In this paper, a fast CU size decision algorithm is presented. The proposed algorithm consists of CU termination and CU skip methods to reduce the redundant computing of inter prediction in HEVC. The offline learning method is used to obtain the statistical parameters. Furthermore, in order to reduce the encoding complexity with negligible loss of encoding efficiency, the MRF-based improvement CU size decision method is presented. The simulation results demonstrate that the overall algorithm can significantly reduce the encoding complexity.

To further enhance the accuracy of CU size decision process, future work can be done by improving the Markov random fields model with neighboring CUs. Furthermore, the parallel computing strategies would be explored to achieve the real-time process for encoder.

Acknowledgment. This research was sponsored by National Natural Science Foundation of China (NSFC, NO. 31170952, 61701297) and the State Oceanic Administration Foundation of China (SOA, NO. 201305026). It was also supported by JSPS KAKENHI Grant Number 17K00157.

REFERENCES

[1] G. J. Sullivan, J. R. Ohm, J. R. Han and T. Wiegand, Overview of the high efficiency video coding (HEVC) standard, *IEEE Trans. Circuits and Systems for Video Technology*, vol.22, no.12, pp.1649-1668, 2012.

[2] S. Ahn, M. Kim and S. Park, Fast decision of CU partitioning based on SAO parameter, motion and PU/TU split information for HEVC, *Picture Coding Symposium (PCS)*, pp.113-116, 2013.

[3] S. Ahn, B. Lee and M. Kim, A novel fast CU encoding scheme based on spatiotemporal encoding parameters for HEVC inter coding, *IEEE Trans. Circuits and Systems for Video Technology*, vol.25, pp.422-435, 2015.

[4] X. Shen, L. Yu and J. Chen, Fast coding size selection for HEVC based on Bayesian decision rule, *Picture Coding Symposium (PCS)*, pp.453-456, 2013.

[5] J. Kim, J. Yang, K. Won and B. Jeon, Early determination of mode decision for HEVC, *Picture Coding Symposium (PCS)*, pp.449-452, 2012.

[6] S. Ahn, B. Lee and M. Kim, Fast zero block detection and early CU termination for HEVC video coding, *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp.1640-1643, 2013.

[7] Z. Pan, S. Kwong, M. T. Sun and J. Lei, Early MERGE mode decision based on motion estimation and hierarchical depth correlation for HEVC, *IEEE Trans. Broadcasting*, vol.60, pp.405-412, 2014.

[8] L. Shen, Z. Liu, X. Zhang, W. Zhao and Z. Zhang, An effective CU size decision method for HEVC encoders, *IEEE Trans. Multimedia*, vol.15, pp.465-470, 2013.
[9] Y. Zhang, H. Wang and Z. Li, Fast coding unit depth decision algorithm for interframe coding in HEVC, *Data Compression Conference (DCC)*, pp.53-62, 2013.
[10] J. Leng, L. Sun, T. Ikenaga and S. Sakaida, Content based hierarchical fast coding unit decision algorithm for HEVC, *Multimedia and Signal Processing (CMSP)*, pp.56-59, 2011.
[11] J. H. Lee, C. S. Park, B. G. Kim, D. J. Sun, S. H. Jung and J. S. Choi, Novel fast PU decision algorithm for the HEVC video standard, *IEEE International Conference on Image Processing (ICIP)*, pp.1982-1985, 2013.
[12] G. Correa, P. Assuncao, L. Agostini and L. A. D. S. Cruz, Coding tree depth estimation for complexity reduction of HEVC, *Data Compression Conference (DCC)*, pp.43-52, 2013.
[13] F. Mu, L. Song, X. Yang and Z. Luo, Fast coding unit depth decision for HEVC, *IEEE International Conference on Multimedia and Expo Workshops*, pp.1-6, 2014.
[14] H. L. Tan, F. Liu, Y. H. Tan and C. Yeo, On fast coding tree block and mode decision for high-efficiency video coding (HEVC), *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp.825-828, 2012.
[15] J. Lee, S. Kim, K. Lim and S. Lee, A fast CU size decision algorithm for HEVC, *IEEE Trans. Circuits and Systems for Video Technology*, vol.25, pp.411-421, 2015.
[16] M. B. Cassa, M. Naccari and F. Pereira, Fast rate distortion optimization for the emerging HEVC standard, *Picture Coding Symposium (PCS)*, pp.493-496, 2012.
[17] J. Vanne, M. Viitanen and D. Hamalainen, Efficient mode decision schemes for HEVC inter prediction, *IEEE Trans. Circuits and Systems for Video Technology*, vol.24, pp.1579-1593, 2014.
[18] Y. Zhang, S. Kwong, X. Wang and H. Yuan, Machine learning based coding unit depth decisions for flexible complexity allocation in high efficiency video coding, *IEEE Trans. Image Processing*, vol.24, pp.2225-2238, 2015.
[19] L. Shen, Z. Zhang and Z. Liu, Adaptive inter-mode decision for HEVC jointly utilizing inter-level and spatio-temporal correlations, *IEEE Trans. Circuits and Systems for Video Technology*, vol.24, pp.1709-1722, 2014.
[20] G. Correa, P. Assuncao, L. Agostini and L. A. da Silva Cruz, Fast HEVC encoding decisions using data mining, *IEEE Trans. Circuits and Systems for Video Technology*, vol.25, pp.660-673, 2015.
[21] G. Correa, P. Assuncao, L. Agostini and L. A. da Silva Cruz, A method for early-splitting of HEVC inter blocks based on decision trees, *The 22nd European Signal Processing Conference (EUSIPCO)*, pp.276-280, 2014.
[22] X. Jiang, T. Song, W. Shi, T. Katayama, T. Shimamoto and L. Wang, Fast coding unit size decision based on probabilistic graphical model in high efficiency video coding inter prediction, *IEICE Trans. Information and Systems*, vol.E99-D, no.11, pp.2836-2839, 2016.
[23] X. Jiang, *Study on Key Technologies for Video Coding Standard Beyond H.265/HEVC*, Ph.D. Thesis, Tokushima University, Tokushima, 2016.
[24] J. Xiong, H. Li, F. Meng, S. Zhu, Q. Wu and B. Zeng, MRF-based fast HEVC inter CU decision with the variance of absolute differences, *IEEE Trans. Multimedia*, vol.16, pp.2141-2153, 2014.
[25] J. Xiong, H. Li, Q. B. Wu and F. Meng, A fast HEVC inter CU selection method based on pyramid motion divergence, *IEEE Trans. Multimedia*, vol.16, pp.559-564, 2014.
[26] D. Chai, S. L. Phung and A. Bouzerdoum, A Bayesian skin/non-skin color classifier using non-parametric density estimation, *Proc. of the 2003 International Symposium on Circuits and Systems (ISCAS)*, pp.464-467, 2003.
[27] J. Yang, J. Kim, K. Won, H. Lee and B. Jeon, Early skip detection for HEVC, *JCTVC-G543*, 2011.
[28] P. Perez, Markov random fields and images, *IRISA*, 2011.
[29] G. Bjontegaard, Calculation of average PSNR differences between RD-curves, *VCEG-M33*, 2001.