Unimodal Model-Based Inter Mode Decision for High Efficiency Video Coding

HUANQIANG ZENG[1], (Senior Member, IEEE), **WENJIE XIANG**[2], **JING CHEN**[1], (Member, IEEE), **CANHUI CAI**[1], (Senior Member, IEEE), **ZHANGKAI NI**[3], (Member, IEEE), **AND KAI-KUANG MA**[4], (Fellow, IEEE)

[1]School of Information Science and Engineering, Huaqiao University, Xiamen 361021, China
[2]Minnan Science and Technology Institute, Fujian Normal University, Quanzhou 362300, China
[3]Department of Computer Science, City University of Hong Kong, Hong Kong
[4]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798

Corresponding author: Jing Chen (jingzi@hqu.edu.cn)

ABSTRACT In this paper, a fast inter mode decision algorithm, called the unimodal model-based inter mode decision (UMIMD), is proposed for the latest video coding standard, the high-efficiency video coding. Through extensive simulations, it has been observed that a unimodal model (i.e., with only one global minimum value) can be established among the size of different prediction unit (PU) modes and their resulted rate-distortion (RD) costs for each quad-tree partitioned coding tree unit (CTU). To guarantee the unimodality and further search the optimal operating point over this function for each CTU, all the PU modes need to be first classified into 11 mode classes according to their sizes. These classes are then properly ordered and sequentially checked according to the class index, from small to large so that the optimal mode can be early identified by checking when the RD cost starts to arise. In addition, an effective instant SKIP mode termination scheme is developed by simply checking the SKIP mode against a pre-determined threshold to further reduce the computational complexity. The extensive simulation results have shown that the proposed UMIMD algorithm is able to individually achieve a significant reduction on computational complexity at the encoder by 61.9% and 64.2% on average while incurring only 1.7% and 2.1% increment on the total Bjontegaard delta bit rate (BDBR) for the low delay and random access test conditions, compared with the exhaustive mode decision in the HEVC. Moreover, the experimental results have further demonstrated that the proposed UMIMD algorithm outperforms multiple state-of-the-art methods.

INDEX TERMS HEVC, coding tree unit, prediction unit, inter mode decision, unimodal model.

I. INTRODUCTION

The latest video coding standard, called the *high efficiency video coding* (HEVC), is a joint effort of two standardization organizations—the *video coding experts group* (VCEG) and the *moving picture experts group* (MPEG) [1]. Compared with its predecessor—the H.264/AVC [2], the HEVC is able to deliver about 50% bit rate reduction while maintaining a

The associate editor coordinating the review of this manuscript and approving it for publication was Byung-Gyu Kim.

similar subjective video quality [3]. However, this is accomplished at a very high cost on computational complexity of HEVC encoder. Unlike the H.264/AVC, the HEVC adopts a much more flexible quad-tree partitioned *coding tree unit* (CTU) structure, which consists of a set of variable-sized *coding units* (CUs), the *prediction units* (PUs), and the *transform units* (TUs). For each CTU, an *exhaustive* evaluation of all available modes under various sizes are performed for identifying the best configuration that will yield the least *rate distortion* (RD) cost. Such so-called *inter* mode decision

performed for each CTU needs to be determined before the actual encoding process of image sequence starts. Inevitably, this inter mode decision process is fairly time consuming. How to reduce the computational complexity while maintaining almost the same coding efficiency has become the central issue of encoding optimization for realizing any practical HEVC encoder.

Since the inception of the H.264/AVC, fast intra and inter mode decision methods have been developed as an effective solution for encoding optimization (e.g., [4]–[8]). However, these mode decision methods cannot be directly applied to the HEVC, since the coding structure of the HEVC is much more flexible and sophisticated than that of the H.264/AVC. It is therefore necessary to develop fast mode decision methods specifically for HEVC and some most recently introduced ones will be highlighted in Section II-C. The main idea of these methods is to only check those more likely CUs/PUs rather than all of them so that the computational complexity can be effectively reduced with acceptable loss in coding performance. Those more likely CUs/PUs for current CTU are generally selected by using various coding information of the spatial/temporal neighboring CTUs, e.g., residual [9], *motion vector* (MV) [10], [11], CU depth [12], [13], optimal mode [14], RD cost [15], texture information [16], all zero blocks [17], sample-adaptive-offset coefficients [18], coded block flag information [18], etc. and various strategies, such as, Bayesian approach [9], [19], [20], conditional random field [21] neighboring prediction [14], [22], adaptive ordering [23], complexity control [24]–[26], linear programming [10], machine learning [27]–[29], etc. Although many efforts has been devoted to HEVC encoder complexity reduction, there is still ample room for further improvement.

For that, an efficient inter mode decision algorithm, called the *unimodal model-based inter mode decision* (UMIMD), is developed for the HEVC in this paper. In our approach, a special relationship among the PU modes and their resulted RD costs has been observed and modeled as a unimodal function. That is, there is only one global minimum value existing in this function. This greatly facilitates the search process on the identification of this minimum point. To guarantee the unimodality and further identify this point, firstly, all the PU modes need to be classified into several *mode classes* according to the PU's size. All these classes need to be properly ordered such that the above-mentioned unimodal property can be yielded for each CTU. To search for the optimal operating point, the RD costs of each pair of two consecutive mode classes are sequentially compared to determine whether the mode class with higher class-index starts to yield a larger RD cost. If so, the global minimum is considered just found (namely, the one with lower class index). In addition, to possibly reduce the computational complexity further, an instant SKIP mode termination scheme is proposed, which is based on an evaluation of the RD cost of the SKIP mode against a pre-determined threshold. Extensive experimental results have shown that the proposed UMIMD

algorithm is able to significantly reduce the computational complexity while maintaining almost the same video coding quality and the total bit rate, compared with the exhaustive mode decision used in the HEVC.

The remaining of this paper is organized as follows. Section II provides a brief overview of the inter mode decision conducted in the HEVC as the background and related fast mode decision methods. Section III describes the proposed fast inter mode decision algorithm for the HEVC in detail. Extensive simulation results are presented and discussed in Section IV. Finally, the conclusion is drawn in Section V.

II. OVERVIEW OF THE INTER MODE DECISION IN THE HEVC AND RELATED WORKS
A. FLEXIBLE CODING STRUCTURE
One of the new features introduced in the HEVC is its highly flexible coding structure, allowing a fixed-sized large block (64×64) to be partitioned in various ways for pursuing better coding efficiency [30]. Analogous to the concept of 16×16 *macroblock* (MB) as used in the previous video coding standard—H.264/AVC, the so-called *coding tree unit* (CTU) is defined in the HEVC as the basic processing block unit with a size of 64×64. Each CTU is subject to be further divided into non-overlapping smaller blocks, called as the *coding units* (CUs) in the HEVC, using the quad-tree partition operation. Note that each CU might be further partitioned to even smaller CUs likewise. Each stage of such quad-tree partition is referred as reaching to the next *depth*, a special parlance as defined in the HEVC. As demonstrated in Fig. 1, the block size of CU ranges from 64×64 (i.e., the CTU) to 8×8. In other words, $CU_{64 \times 64}$ corresponds to $depth = 0$; hence, $CU_{32 \times 32}$ is at $depth = 1$, $CU_{16 \times 16}$ is at $depth = 2$, and $CU_{8 \times 8}$ is at $depth = 3$.

For each $2N \times 2N$ CU in any depth, it has multiple combinations of various-sized partitioned rectangular blocks, and each combination is called a *prediction unit* (PU) and defined as an unique *mode* in the HEVC. This set of modes as defined by the HEVC is subject to conduct inter prediction or intra prediction for each CU individually to determine which one is the best mode for that CU (Refer to Fig. 2). In addition to the SKIP mode, these modes include $2N \times 2N$, $2N \times N$, $N \times 2N$, $N \times N$, $nL \times 2N$, $nR \times 2N$, $2N \times nU$, and $2N \times nD$, to effectively employ the temporal correlation inherited in video. For performing *intra* prediction, there are only two types of PU modes: $2N \times 2N$ and $N \times N$, where the latter one is only available for the 8×8 CU (i.e., $N = 4$).

Lastly, there are three *transform unit* (TU) modes have been introduced in the HEVC for conducting DCT-based transform and quantization of prediction residuals. Three TU modes are with different block sizes, including $2N \times 2N$, $N \times N$, and $N/2 \times N/2$. In summary, Fig. 2 shows the available PU modes and TU modes for a $2N \times 2N$ CU.

B. EXHAUSTIVE MODE DECISION
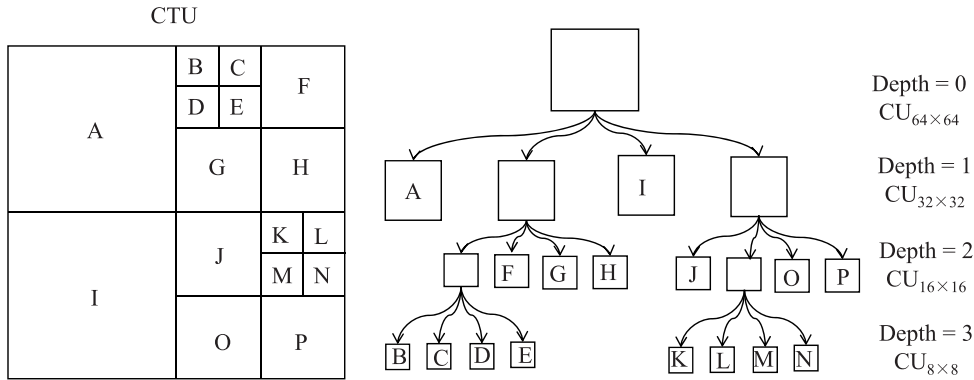Although the above-described flexible coding structure is able to yield a much higher coding efficiency (compared to

FIGURE 1. An example of the *coding tree unit* (CTU), which is the basic processing unit as practiced in the HEVC. Each CTU consists of a set of *coding units* (CUs) that are recursively generated using quad-tree partition process. Note that the block size of the CU is variable.
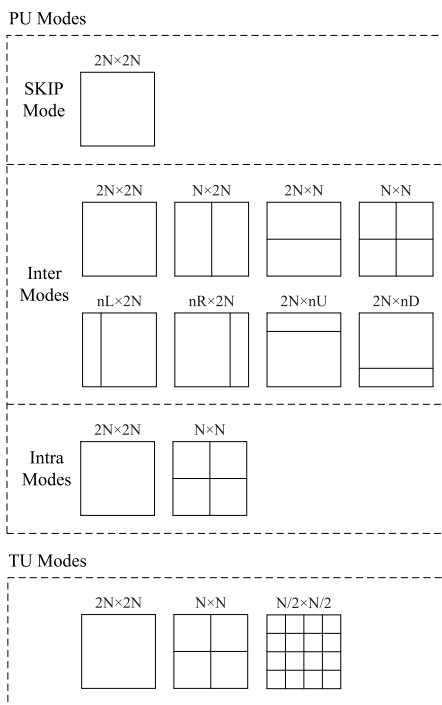

FIGURE 2. PU modes and TU modes for a 2*N* × 2*N* CU.

identify which one has the least RD cost that corresponds to the optimal mode combination. The Lagrangian RDO function is defined as

$$J = SSE(s, c|QP) + \lambda \cdot R(s, c|QP), \qquad (1)$$

where QP is the quantization parameter, λ is the Lagrangian multiplier related to QP, R represents the number of bits required, SSE stands for the *sum of squared error* calculated between the original luma block (i.e., s) and its reconstructed block (i.e., c); that is,

$$SSE = \sum_{(x,y) \in s,c} |s(x, y) - c(x, y)|^2. \qquad (2)$$

Although the above-described exhaustive mode decision strategy is able to find the optimal mode combination for ensuring the high coding efficiency, however the involved computational complexity is extremely high as expected. Therefore, it is very essential to develop a fast mode decision method for the HEVC, aiming to significantly reduce the computational complexity without causing the degradation on coding efficiency. In what follows, we shall highlight several state-of-the-art fast inter mode decision methods for HEVC before presenting ours.

C. LITERATURE REVIEW

Multiple fast mode decision methods for the HEVC can be found in the literature. Shen *et al.* [9] developed a fast CU decision method by exploiting the Baysian decision rule, where the conditional probability is computed based on the evaluation of the *sum of absolute transformed difference* (SATD), the magnitude of MV, and the RD cost. Zhao *et al.* [10] presented a flexible CU decision method by exploiting the mode mapping scheme, together with the use of an optimal stopping criterion, to reduce the computational load. Mallikarachchi *et al.* [11] presented a content-adaptive feature-based CU decision method. In this method, two CU split likelihood models are trained to model the CU split and non-split decisions, and a motion vector reuse mechanism is developed to further expedite the motion estimation process.

that of the previous video coding standard—H.264/AVC), the key concern is its involved computational complexity on the determination of the optimal mode combination—namely, the best combination of the optimal CU partitions, PU, and TU modes among all available modes. For that, the HEVC reference software (i.e., HM) exploits the so-called *exhaustive mode decision* strategy to assure that the optimal one will be identified at the end of mode decision process.

In the HM, the Lagrangian *rate distortion optimization* (RDO) function [31] is exploited as the mode decision criterion. That is, all the CU, PU and TU modes need to be exhaustively evaluated by computing the RD cost resulted from each mode. These computed costs will be compared to

Shen *et al.* [12] presented a fast mode decision algorithm that consists of two schemes—adaptive depth range determination and early termination of unnecessary *motion estimation* (ME) process that might be encountered on small-sized CUs. Tai *et al.* [13] avoided CU partition procedure and truncated PU prediction by investigating the depth correlation between the current CU and the collocated CU, and analyzing the RD cost in different cases. Shen *et al.* [14] proposed a hybrid mode decision method, in which all the candidate modes are firstly classified into three mode classes. Among them, only one will be chosen to conduct further evaluations based on the mode complexity of the current CU, which is measured according to the coding information of its neighboring and temporally co-located CUs. Two early termination schemes are further incorporated to early select the SKIP/PU mode as the optimal mode of the current CU by using the mode and RD cost correlations between the current CU and its spatially and/or temporally adjacent CUs. Vanne *et al.* [15] addressed the complexity overhead of the symmetric motion partition and asymmetric motion partition in the HEVC by analyzing the rate-distortion complexity of the HEVC inter prediction as a function of different block partition structures. Goswami *et al.* [16] proposed a low complexity coding method for HEVC through texture analysis. Pan *et al.* [17] proposed an early MERGE mode decision method by utilizing ME information and all zero blocks. Ahn *et al.* [18] proposed a fast CU encoding scheme by making full use of the spatial and temporal encoding parameters produced by the HEVC encoder. In this method, the spatial encoding parameters, sample-adaptive-offset coefficients, and the temporal encoding parameters, including MV, TU sizes, and the *coded block flag* information, are used to individually estimate the texture characteristic and temporal complexity of each CU, which are then utilized to conduct early SKIP mode detection and fast CU split for significantly reducing the computational complexity. Lee *et al.* [19] proposed a method involving three stages of early terminations, in which three thresholds are designed by using Bayes' rule with a complexity factor. Goswami and Kim [20] developed a fast scheme for HEVC, in which the Baysian classifier is designed by treating SKIP detection and CU termination as two-class decision problem and trained by using Markov chain Monte Carlo model. Zhang *et al.* [21] developed a two-stage fast inter CU decision method by exploring CU classification based on Bayesian rule, and early CU skipping/pruning scheme based on conditional random fields. Duan *et al.* [22] presented an adaptive quad-tree depth range prediction mechanism based on the depth correlation within the neighboring region. Jung and Park [23] explored an adaptive ordering of various modes and prediction of the RD cost and bit cost to reduce the computational complexity of HEVC. Jimenez-Moreno *et al.* [24] developed an effective complexity control algorithm in a hierarchical manner. This method estimated all the parameters based on fly so that the encoder behavior can be adapted to the video content, the encoding configuration, and the target complexity over time. Zhang *et al.* [25]

presented a novel CTU-level complexity control method, in which the complexity of each CTU is first estimated by a statistical model and is then used to to optimize its depth range. Also, a complexity control scheme, including complexity estimation, complexity allocation, and PU adaption is designed for HEVC intra coding [26]. Xiong *et al.* [27] formed the *pyramid motion divergence* (PMD) by computing the optical flows on the down-sampled frames to obtain the motion information, followed by employing a k-nearest neighborhood like algorithm to determine whether the current CU needs to be split or not. Zhang *et al.* [28] presented a machine learning-based fast CU decision method. Firstly, the CU decision process is modeled as a three-level hierarchical decision problem. Then, several techniques, including an improved CU decision structure, a three-output classifier, and a sophisticated RD-complexity model, are developed to achieve a good trade-off between the complexity reduction and the RD maintenance. Zhu *et al.* [29] formulated the CU decision process as a cascaded multi-level classification task and solved it with the fuzzy support vector machine.

III. PROPOSED UNIMODAL MODEL-BASED INTER MODE DECISION (UMIMD) ALGORITHM FOR HEVC

A. UNIMODAL MODEL

Note that the above-mentioned prediction modes as exploited in the HEVC were designed for better adapting to different video contents for yielding high coding efficiency. To be more specific, the SKIP mode is more proper for the smooth region, the larger block-sized modes are more beneficial to those regions with more homogeneous content and with slow motion, the smaller block-sized modes are suitable for those regions involving fast motion, and the *intra* modes are most appropriate for the inhomogeneous regions with complex texture. Intuitively, it would be logical to skip the checking process for the smooth CTU with motionless or fairly slow motion. In other words, according to the evaluation of the smoothness and motion activity of the CTU, the checking process of those unlikely modes can be skipped to reduce the computational complexity of *inter* mode decision in HEVC. Further studies have shown that the smoothness and motion activity of the CTU can be accurately reflected by its RD cost value. That is, if a CTU is not smooth, not only the difference value (i.e., *SSE*) between the original pixels and the reconstructed pixels will be larger but also more bits will be required for coding the residual and mode information. In this case, a larger RD cost value as computed in (1) will be resulted. Hence, this relationship can be utilized on the design of fast mode decision algorithm as proposed in this paper.

Considering a PU is the basic processing unit for inter/intra prediction, the relationship between each PU mode and its yielded RD cost value is important to be investigated as follow. First, note that the total bits denoted as R in (1) includes the bits required to code the prediction residual (denoted as R_r) and that for coding the MVs and mode information (denoted as R_m); thus, the RD cost function in (1) can be

expressed as

$$J = SSE + \lambda \cdot (R_r + R_m). \qquad (3)$$

Note that when the size of chosen PU mode is small, the prediction tends to be more accurate. This leads to less normalized RD cost to PU size and a smaller residual, which requires less R_r bits to represent the residual information. However, all these are achieved at the expense of more prediction vectors need to be encoded. For example, in a given CU, only one MV needs to be transmitted for inter $2N \times 2N$ mode while four MVs are required to be represented and transmitted for inter $N \times N$ mode. Hence, this will inevitably cost more bits on R_m due to the increment of mode information and thus increase the RD cost value. In other words, when the size of PU mode becomes smaller and smaller, the normalized bits to PU size consumed on R_r will be monotonically *decreasing*, while the normalized bits to PU size spent on R_m will be monotonically *increasing*.

TABLE 1. Mode classes and their involved PU modes.

Class n	0	1	2
Mode	SKIP_64×64	64×64	64×32, 32×64
Class n	3	4	5
Mode	SKIP_32×32	32×32	32×16, 16×32
Class n	6	7	8
Mode	SKIP_16×16	16×16	16×8, 8×16
Class n	9	10	
Mode	SKIP_8×8	P8×8, Intra	

Inspired by the above-mentioned trend, how to classify and group all the available HEVC's modes in an effective way is a crucial step for the mode decision process. For that, all the HEVC's PU modes are categorized into 11 *mode classes* in this paper and summarized in Table 1. This classification and grouping is according to the sizes of various PU modes. Note that although the SKIP mode and the inter $2N \times 2N$ (where $N = 4, 8, 16,$ and 32) have the same size (e.g., SKIP_64×64 and 64×64), but they are more suitable for different kinds of regions. For example, the SKIP mode is beneficial to homogeneous regions and thus classified as a special mode class. Besides, the intra modes are grouped together P8 × 8 (i.e., 8×8, 8×4, 4×8, and 4×4), since they are all more suitable on coding those regions with highly textures. Moreover, considering the fact that the asymmetric PU modes has lower possibility to be selected as the optimal mode, therefore, only the symmetric PU modes are considered in our work for conducting the mode decision. For ease of development and presentation, the class index n used in Table 1 is assigned to the PU modes according to their sizes, from large to small. We further denote the resulted bits of R_r and R_m incurred at class n as $R_r(n)$ and $R_m(n)$, respectively. When the size of PU mode is small (i.e., large n), less $R_r(n)$, but more $R_m(n)$, will be resulted. As a result, when the mode class changes from n to $n+1$, it becomes difficult to evaluate the net contribution or gain based on (3). To analyze this, let $J(n)$ be the minimum RD cost resulted from the mode class n, $\triangle D_r(n)$ be the reduced RD cost resulted from the improvement of prediction accuracy, and $\triangle D_m(n)$ be the increased RD cost for encoding more mode information. Consequently, the relationship between the RD cost of the mode classes n and $n+1$ can be described as

$$J(n+1) = J(n) + \triangle D_r(n) + \triangle D_m(n). \qquad (4)$$

Clearly, if $|\triangle D_r(n)| > |\triangle D_m(n)|$, the usage of a smaller block-sized PU mode (i.e., using mode class $n + 1$ instead of n) will reduce the RD cost and thus be beneficial to improve the coding efficiency. On the contrary, if $|\triangle D_r(n)| < |\triangle D_m(n)|$, the usage of smaller block-sized PU mode will increase the RD cost, which is obviously undesirable let alone it has incurred extra computational complexity. Further investigation has found out that when $|\triangle D_r(n)| < |\triangle D_m(n)|$, the residual has become quite small, which is hard to further reduced by using a smaller block-sized PU mode. In this situation, if the size of PU mode is further reduced (i.e., with a larger mode class number n), the change of $|\triangle D_r(n)|$ value tends to become ignorable and with very limited coding gain to be yielded. On the other hand, the mode information will be substantially increased by the geometric progression, leading a much higher increment of $|\triangle D_m(n)|$. Therefore, $|\triangle D_r(n+1)| < |\triangle R_m(n+1)|$ will be usually satisfied when $|\triangle D_r(n)| < |\triangle D_m(n)|$. Hence, a special relationship among the PU modes and their resulted RD costs can be observed and modeled as a *unimodal* model (i.e., with only one global minimum value incurred on the RD cost function).

TABLE 2. Sequences for training and modeling.

Sequences	Resolutions	Sequences	Resolutions
BasketballPass	416×240	PartyScene	832×480
BlowingBubbles	416×240	Kimono	1920×1080
BasketballDrill	832×480	BQTerrace	1920×1080

To verify this unimodal model, extensive experiments have been conducted by applying the exhaustive mode decision in the HEVC to the video sequences as listed in Table 2. Fig. 3 illustrates the relationship between the RD cost of different block-sized PU modes $J(n)$ and mode class n under various optimal PU modes. Note that $J(n)$ for each mode class n is computed and normalized by its PU size when various PU modes are exploited as the optimal mode for the current CU. Take Fig. 3(d) as an example. When the SKIP_32 × 32 (i.e., the mode class 3) is the optimal mode of $CU_{32 \times 32}$, one can see that the global minimum point over the $J(n)$ function is incurred at $n = 3$. Same interpretation can be applied to other sub-graph in Fig. 3 when their global minimum (i.e., the optimal mode) are incurred at various class n values. It is important to note that no matter which optimal PU mode is identified, the $J(n)$ always has such unimodal property (i.e., with only one global minimum) in terms of the size of PU mode (i.e., the mode class n). By utilizing this unique
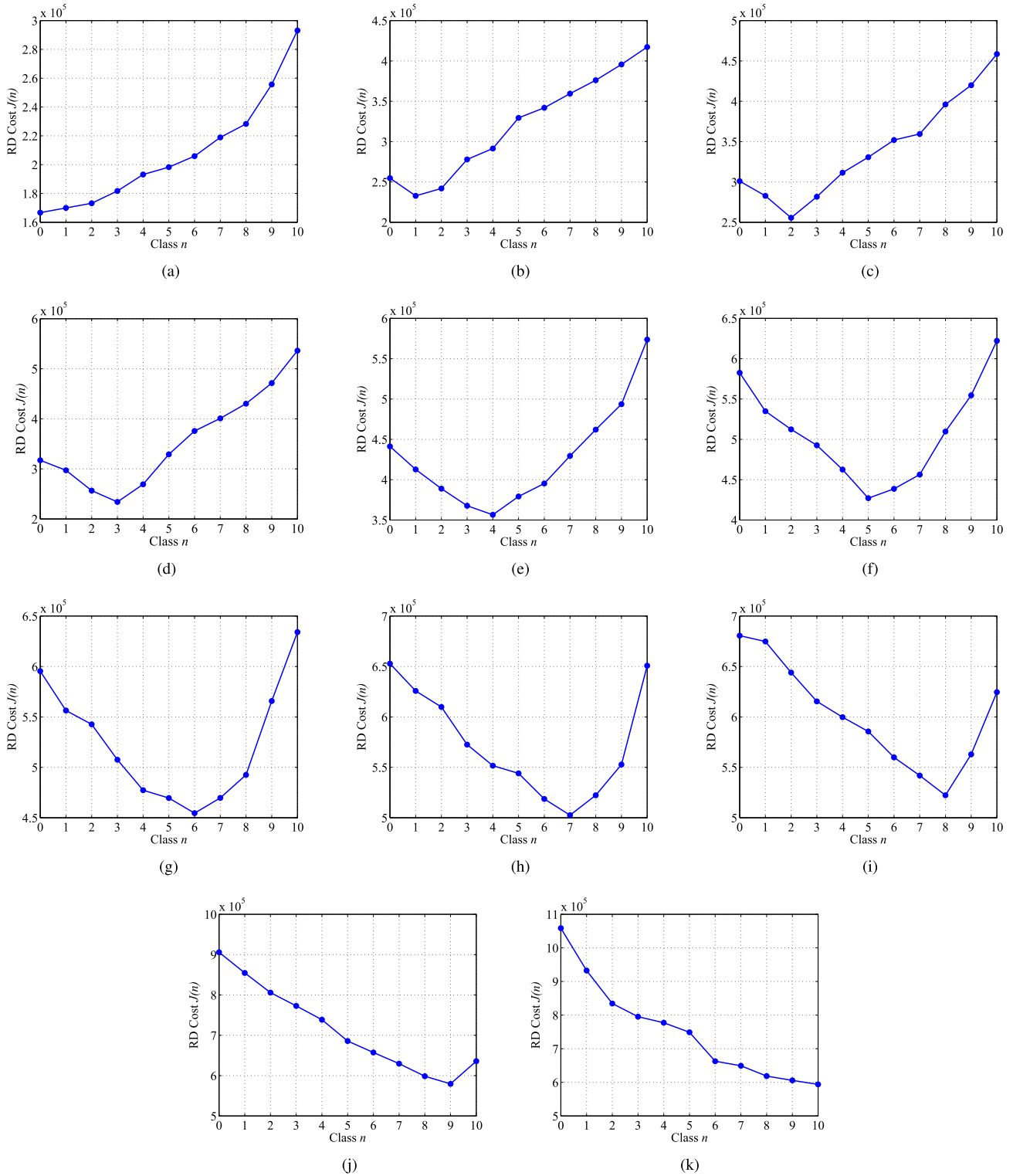
FIGURE 3. The relationship between the RD cost and the size of PU modes under various optimal PU modes. (a) Optimal mode belongs to Class 0. (b) Optimal mode belongs to Class 1. (c) Optimal mode belongs to Class 2. (d) Optimal mode belongs to Class 3. (e) Optimal mode belongs to Class 4. (f) Optimal mode belongs to Class 5. (g) Optimal mode belongs to Class 6. (h) Optimal mode belongs to Class 7. (i) Optimal mode belongs to Class 8. (j) Optimal mode belongs to Class 9. (k) Optimal mode belongs to Class 10.

property, the principle of the proposed UMIMD algorithm is described as follows.

The mode classes as listed in Table 1 can be sequentially checked according to the class index *n*, either from small to

TABLE 3. Distribution of the optimal mode class in the HEVC (%).

Sequences		BasketballPass	BlowingBubbles	BasketballDrill	PartyScene	Kimono	BQTreeace	Average
Class n	Mode							
0	SKIP_64×64	18.40	8.55	26.05	19.51	26.96	47.84	24.55
1	64×64	4.96	3.73	14.64	3.33	14.21	8.21	8.18
2	64×32/32×64	2.85	4.06	4.12	3.05	5.29	4.89	4.04
3	SKIP_32×32	26.47	27.51	18.80	23.24	17.94	16.83	21.80
4	32×32	5.32	8.65	10.39	9.76	16.58	5.59	9.38
5	32×16/16×32	1.59	3.85	2.06	4.74	2.75	1.62	2.77
6	SKIP_16×16	18.83	18.07	8.58	12.54	6.36	6.41	11.80
7	16×16	9.28	12.00	8.77	11.53	6.93	3.59	8.68
8	16×8/8×16	1.57	2.07	1.11	2.63	0.42	0.63	1.41
9	SKIP_8×8	5.69	4.46	3.00	6.09	1.62	2.68	3.92
10	P8×8/Intra	5.04	7.05	2.48	3.58	0.94	1.71	3.47

large or from large to small, so that the optimal mode can be easily identified by checking when the RD cost starts to arise. When this happens, it means that the optimal mode (i.e., the global minimum point) has just incurred in the previous mode class checking. In this work, considering that the larger PU mode is more likely to be the optimal one while incurs lower computational complexity than the smaller PU mode, the mode classes as listed in Table 1 will be sequentially checked according to the class index n, from small to large. In other words, the RD costs of two consecutive mode classes will be sequentially compared to determine whether the mode class with higher class-index starts to yield a larger RD cost; i.e., checking whether $J(n + 1) > J(n)$. If so, the global minimum is considered found at class index n, and the checking process of the remaining modes can be skipped, since the optimal mode has been identified. Consequently,the computational load can be significantly saved.

B. INSTANT SKIP MODE TERMINATION SCHEME

The above-described fast mode decision scheme using the *unimodal model* (denoted as UM) needs to compute the RD cost of the next mode class $n + 1$ in order to determine whether the RD cost $J(n + 1)$ is greater than $J(n)$. Obviously, this consumes extra computational complexity. To further reduce unnecessary computation of UM algorithm, an instant-termination scheme is desired to instantly determine whether the mode class under consideration is the optimal one. For that, extensive simulation experiments have been conducted to study the distribution of the *optimal* mode class presented in the HEVC by exploiting the exhaustive mode decision based on the video sequences as listed in Table 2 under the common test conditions [32], and the results are shown in Table 3.

This investigation clearly indicates that for the variable block-sized CUs, the mode classes 0, 3, 6, 9 (i.e., the SKIP modes with variable block sizes) have higher possibility to be the optimal one, especially for those sequences with motionless or slow-motion content. This is because large homogeneous, motionless or slow motion scenes are often encountered in the real-world video sequences while the

SKIP mode is more suitable for coding such kind of scenes. On the other hand, the checking process of the SKIP mode has lower computational complexity, since the SKIP mode does not require to conduct ME operation; therefore, neither the MV/reference indices nor the texture residuals will be incurred for encoding and transmission (In fact, the MV is derived from the MVs of the spatial/temporal neighboring CUs). Therefore, if the SKIP mode can be determined instantly (i.e., the RD cost of the next mode class needs not to be computed), the computational complexity will be substantially reduced again. Since the motion activity of the CU is highly related to its RD cost value, an instant SKIP mode termination scheme is developed by evaluating whether the RD cost computed at the SKIP mode is below an empirically pre-determined threshold. More specifically, for the variable block-sized CUs represented by different depths within the current CTU (See Fig. 1), if the RD cost of the SKIP mode computed at the current depth is small enough (that is, below a threshold T corresponding to the current depth), then the current CU is quite likely smooth and/or with fairly slow motion. In this case, the SKIP mode is thus selected as the optimal mode, and the mode decision process proceeds to the next CU.

Now, the key question is how to find a robust value of the threshold for each depth. For the CTU as shown in Fig. 1, four thresholds, T_0, T_1, T_2, and T_3, are proposed in our mode decision algorithm for variable block-sized CUs that correspond to the depth levels 0, 1, 2, and 3, respectively. For example, to determine T_0 for $CU_{64\times64}$ (i.e., depth = 0), all the $CU_{64\times64}$ from a set of commonly-used test sequences listed in Table 2 are employed to determine the most reliable value of the threshold T_0—with a goal of achieving 95% degree of confidence. In other words, if the RD cost value of the SKIP mode computed at the $CU_{64\times64}$ (i.e., for the mode class 0, SKIP_64 × 64) is smaller than T_0, the SKIP_64 × 64 mode is assigned to the current $CU_{64\times64}$ as its optimal mode, and the probability that this SKIP_64 × 64 mode is indeed the optimal mode according to the exhaustive mode decision is not less than 0.95. The thresholds T_1, T_2, and T_3 are determined in a similar manner as performed for the T_0. Further consider that

TABLE 4. Empirically-determined threshold values, T_1, T_2, T_3, and T_4, with 95% degree of confidence under different *QP* values.

QP	12	17	22	27	32	37	42
T_0	2100	8700	25100	97008	230300	527400	1402100
T_1	850	2640	9523	19640	55690	129200	417310
T_2	230	660	1880	5370	15350	43860	103140
T_3	61	161	426	1129	2994	8036	21040

different *QP*s could yield various degree of quantization of the video data, these thresholds are dependent on the *QP*, and their values for various *QP* are shown in Table 4. It has been observed that, approximately, there is an exponential relationship existing between the threshold values T_d and various *QP* values. To mathematically model this relationship, a Matlab curve fitting tool, `cftool`, is applied to approximate an exponential function with the format as: $T_d = ae^{b \times QP}$, where parameters a and b are to be determined based on the data points in Table 4. The resulted formulas are

$$T_0 = 512.8026\, e^{0.1884 \times QP}, \tag{5}$$
$$T_1 = 39.3489\, e^{0.2206 \times QP}, \tag{6}$$
$$T_2 = 49.6989\, e^{0.1820 \times QP}, \tag{7}$$
$$T_3 = 6.1716\, e^{0.1937 \times QP}. \tag{8}$$

C. ALGORITHM DESCRIPTION OF THE PROPOSED UMIMD ALGORITHM

In summary, the proposed UMIMD algorithm for the current CTU can be described as follows:

1) For the $CU_{64 \times 64}$, calculate the RD cost of the SKIP_64 $\times$ 64 mode (i.e., the Class 0), $J(0)$.

2) If $J(0) < T_0$, then the SKIP_64 $\times$ 64 mode will be chosen as the optimal mode, and go to Step (7).

3) Calculate the RD cost of the inter 64×64 (i.e., the Class 1), $J(1)$. If $J(1) > J(0)$, choose SKIP_64 $\times$ 64 (i.e., Class 0) as the optimal mode, and go to Step (7).

4) Calculate the RD cost of the inter 64×32 and the inter 32×64 (both from the Class 2), respectively. Select the one with the smaller RD cost as the mode 2 and set the corresponding RD cost as $J(2)$. If $J(2) > J(1)$, choose the inter 64×64 (i.e., the Class 1) as the optimal mode, and go to Step (7).

5) Divide the $CU_{64 \times 64}$ into four $CU_{32 \times 32}$ units. Calculate the RD cost of the SKIP_32 $\times$ 32 (i.e., the Class 3) for each $CU_{32 \times 32}$ and denote them as $J_i(3)$ (where $i = 0, 1, 2$, and 3), respectively. The RD cost of the $CU_{64 \times 64}$ computed at the Class 3 is calculated by summing them together; i.e., $J(3) = \sum_{i=0}^{3} J_i(3)$. If $J(3) > J(2)$, choose the mode 2 as the optimal mode, and go to Step (7).

6) If all the $CU_{32 \times 32}$ are checked, go to Step (7); otherwise, go to Step (6.1).

   6.1) For the *i*-th $CU_{32 \times 32}$, if $J_i(3) < T_1$, then the SKIP_32 $\times$ 32 mode will be chosen as the optimal mode, and go to Step (6).

6.2) Calculate the RD cost of the inter 32×32 (i.e., the Class 4), $J_i(4)$. If $J_i(4) > J_i(3)$, choose the SKIP_32 $\times$ 32 as the optimal mode, and go to Step (6).

6.3) Calculate the RD cost of the inter 32×16 and the inter 16×32 (both are from the Class 5), respectively. Set the one with the smaller RD cost as the mode 5 and the corresponding RD cost as $J_i(5)$. If $J_i(5) > J_i(4)$, choose the inter 32×32 as the optimal mode, and go to Step (6).

6.4) Divide the current (i.e., *i*-th) $CU_{32 \times 32}$ into four $CU_{16 \times 16}$ units. Calculate the RD cost of SKIP_16 $\times$ 16 (i.e., the Class 6) for each $CU_{16 \times 16}$ and denote them as $J_{ij}(6)$ (for $j = 0, 1, 2$, and 3), respectively. The RD cost of the current $CU_{32 \times 32}$ computed at the Class 6 is calculated by summing the computed values together; i.e., $J_i(6) = \sum_{j=0}^{3} J_{ij}(6)$. If $J_i(6) > J_i(5)$, choose the mode 5 as the optimal mode, and go to Step (6).

6.5) If all the $CU_{16 \times 16}$ are checked, go to Step (6); otherwise, go to Step (6.5.1).

   6.5.1) For the (i, j)-th $CU_{16 \times 16}$, if $J_{ij}(6) < T_2$, then the SKIP_16 $\times$ 16 mode will be chosen as the optimal mode, and go to Step (6.5).

   6.5.2) Calculate the RD cost of the inter 16×16 (i.e., the Class 7), $J_{ij}(7)$. If $J_{ij}(7) > J_{ij}(6)$, choose the SKIP_16 $\times$ 16 as the optimal mode, and go to Step (6.5).

   6.5.3) Calculate the RD cost of the inter 16×8 and the inter 8×16 (both are from the Class 8), respectively. Set the one with the smaller RD cost as the mode 8 and the corresponding RD cost as $J_{ij}(8)$. If $J_{ij}(8) > J_{ij}(7)$, choose the inter 16×16 as the optimal mode, and go to Step (6.5).

   6.5.4) Divide the current (i.e., (i, j)-th) $CU_{16 \times 16}$ into four $CU_{8 \times 8}$ units. Calculate the RD cost of the SKIP_8 $\times$ 8 (i.e., Class 9) for each $CU_{8 \times 8}$ and denote them as $J_{ijk}(9)$ (for $k = 0, 1, 2$, and 3), respectively. The RD cost of the current $CU_{16 \times 16}$ computed at the Class 9 is calculated by summing them together; i.e., $J_{ij}(9) = \sum_{k=0}^{3} J_{ijk}(9)$. If $J_{ij}(9) > J_{ij}(8)$, choose the mode 8 as the optimal mode, and go to Step (6.5).

   6.5.5) If all the $CU_{8 \times 8}$ are checked, go to Step (6.5); otherwise, go to Step (6.5.5.1).

      6.5.5.1) For the (i, j, k)-th $CU_{8 \times 8}$, if $J_{ijk}(9) < T_3$, then the SKIP_8 $\times$ 8 mode will be chosen as the optimal mode, and go to Step (6.5.5).

      6.5.5.2) Calculate the RD cost of the P8 $\times$ 8 and the intra-prediction modes (i.e., the Class 10), respectively. Set the one with the smaller RD cost as the mode 10 and

TABLE 5. Performance comparison between proposed UMIMD algorithm and recent works under the LD condition.

Resolution	Sequences	Shen et al. [12]		Shen et al. [14]		Ahn et al. [18]		Zhang et al. [28]		Zhang et al. [21]		Proposed	
		BDBR	TS	BDBR	TS	BDBR	TS	BDBR	TS	BDBR	TS	BDBR	TS
Class A 2560×1600	Traffic	2.0	41.3	1.0	54.7	2.0	41.7	2.0	56.1	0.9	48.4	2.3	70.8
	PeopleOnStreet	2.2	19.9	0.4	32.5	1.5	28.0	1.2	44.4	1.0	41.3	1.0	56.4
Class B 1920×1080	Kimonol	0.3	27.8	0.8	38.0	0.8	47.9	2.0	56.1	0.9	44.7	1.8	62.1
	ParkScene	0.9	28.9	0.8	40.5	1.1	48.3	1.5	47.8	0.4	40.9	2.2	67.7
	Cactus	3.1	36.4	0.5	43.5	2.2	48.0	1.8	48.4	0.5	42.6	2.5	63.6
	BasketballDrive	1.0	25.7	0.8	42.3	1.2	42.6	2.2	51.0	1.4	49.6	1.8	67.5
	BQTerrace	1.8	32.2	1.3	42.7	0.3	47.5	2.1	45.1	2.0	47.2	2.0	63.9
Class C 832×480	BasketballDrill	1.3	38.5	1.9	44.0	2.0	37.0	1.7	47.6	0.7	37.7	1.0	53.8
	BQMall	2.2	25.3	2.0	43.5	1.2	40.9	1.6	46.0	1.0	38.7	1.5	54.6
	PartyScene	3.3	25.1	1.0	40.0	0.4	33.0	1.3	36.6	0.4	24.5	1.3	51.3
	RaceHorses	1.1	14.0	1.0	34.1	1.1	26.1	2.0	47.5	1.0	43.8	1.0	48.5
Class D 416×240	BasketballPass	1.6	14.8	1.5	35.4	1.2	27.7	1.2	36.6	0.5	42.5	1.2	52.8
	BQSquare	0.4	44.3	0.4	36.0	0.2	38.8	0.8	30.7	0.1	33.3	1.8	57.1
	Blowingbubbles	2.4	22.0	2.2	39.7	0.3	31.5	0.9	28.8	0.4	29.2	1.6	51.1
	RaceHorses	1.0	12.9	0.7	30.8	0.7	21.2	1.2	33.4	0.7	39.1	1.0	49.9
Class E 1280×720	FourPeople	2.4	47.8	1.1	64.0	1.6	65.6	2.8	66.4	1.9	70.9	1.6	77.3
	Johnny	2.0	56.0	1.3	66.9	-0.3	73.9	2.5	70.9	2.1	76.0	2.3	86.0
	KristenAndSara	2.8	48.5	1.4	60.9	0.5	69.6	2.3	68.2	2.7	74.8	2.0	79.8
Average		1.8	31.2	1.1	43.9	1.0	42.7	1.7	47.9	1.0	45.8	1.7	61.9

the corresponding RD cost as $J_{ijk}(10)$. If $J_{ijk}(10) > J_{ijk}(9)$, choose the SKIP_8 × 8 as the optimal mode; otherwise, choose the mode 10 as the optimal mode. Go to Step (6.5.5).

7) Proceed to the mode decision of the next CTU and repeat the above-described steps.

IV. EXPERIMENTAL RESULTS

To evaluate the performance, the proposed fast mode decision, UMIMD, algorithm has been incorporated into the HEVC reference software (i.e., HM16.7) [33] and tested on multiple test sequences, covering a wide range of motion activities and resolutions. All the experiments are conducted on a computer with 3.6 GHz Intel i7 Core 4 processors, 8 GB memory, and Windows 7 operating system. The test conditions are set as follows:

1) Each test sequence is encoded using *low delay* (LD) and *random access* (RA) configuration defined in the common test conditions [32].
2) The maximum and minimum CU sizes are 64 × 64 and 8 × 8, respectively.
3) Four QPs, 22, 27, 32 and 37, respectively, are considered in our experiments.
4) The RDO is enabled, and the values of other parameters are set at their default, respectively.

In this work, the performance evaluation of the proposed UMIMD algorithm and five state-of-the-art methods (i.e., Shen *et al.* [12], Shen *et al.* [14], Ahn *et al.* [18], Zhang *et al.* [28], and Zhang *et al.* [21]) are compared under the LD and RA conditions [32], respectively. In our experiments, the objective measurement of the performance gain of each method is referenced to that of the exhaustive mode decision in the HEVC, measured in terms of the following two performance indexes: 1) the *Bjontegaard delta bit rate* (BDBR) [34] measures the *averaged* bit rate changes; 2) *TS* measures the *averaged* complexity reduction over four *QP*s, which can be computed as

$$TS = \frac{1}{4} \sum_{i=1}^{4} \frac{T_{HEVC}(QP_i) - T_{fast}(QP_i)}{T_{HEVC}(QP_i)} \times 100\% \quad (9)$$

where $T_{HEVC}(QP_i)$ and $T_{fast}(QP_i)$ are the encoding time resulted from the exhaustive mode decision in the HEVC and the corresponding fast mode decision methods under four QP values as above-mentioned, respectively.

Tables 5 and 6 show the performance comparison under the LD and RA conditions, respectively. From the results as shown in Tables 5 and 6, it can be seen that compared with the outcomes resulted from the exhaustive mode decision in the HEVC, the proposed UMIMD algorithm constantly achieves, on average, 61.9% and 64.2% time saving, with only 1.7% and 2.1% increment in the BDBR bit rate for the LD and RA test conditions, respectively. Moreover, compared with Shen *et al.* [12], Shen *et al.* [14], Ahn *et al.* [18], Zhang *et al.* [28], and Zhang *et al.* [21], for the LD test condition, 30.7%, 18.0%, 19.2%, 14.0%, and 16.1% higher computational complexity reduction with similar coding efficiency measured in terms of BDBR bit rate changes are achieved by the proposed UMIMD algorithm, respectively; for the RA test conditions, the proposed UMIMD algorithm also yields 31.6%, 19.1%, 14.6%, 12.3%, and 9.3% more time saving while maintaining the similar coding performance, respectively. All these reveal that the proposed UMIMD algorithm consistently performs well and outperforms

TABLE 6. Performance comparison between proposed UMIMD algorithm and recent works under the RA condition.

Resolution	Sequences	Shen *et al.* [12]		Shen *et al.* [14]		Ahn *et al.* [18]		Zhang *et al.* [28]		Zhang *et al.* [21]		Proposed	
		BDBR	TS	BDBR	TS	BDBR	TS	BDBR	TS	BDBR	TS	BDBR	TS
Class A 2560×1600	Traffic	2.1	44.1	1.1	60.5	0.8	61.6	1.9	61.7	1.4	62.8	3.2	71.7
	PeopleOnStreet	4.0	23.3	0.2	42.5	0.9	26.9	1.5	48.3	1.0	47.7	2.0	45.6
Class B 1920×1080	Kimono1	0.4	31.2	1.0	47.3	1.3	58.2	4.0	56.9	1.1	54.6	2.4	66.1
	ParkScene	1.0	33.7	0.9	45.2	1.2	52.6	1.6	54.6	0.6	53.4	2.5	69.1
	Cactus	3.2	40.6	1.0	42.1	2.8	56.8	2.2	52.7	0.8	51.6	3.0	64.1
	BasketballDrive	1.4	28.7	1.0	41.8	2.0	50.9	4.3	54.5	2.0	58.7	1.7	60.0
	BQTerrace	1.2	35.7	1.1	49.7	1.6	54.5	1.8	51.3	1.8	57.0	2.1	72.0
Class C 832×480	BasketballDrill	5.2	34.1	2.1	39.8	1.9	45.2	1.8	52.0	1.1	51.8	1.3	54.1
	BQMall	2.9	28.8	1.6	40.0	2.2	48.6	2.3	49.8	1.6	50.9	2.8	56.1
	PartyScene	2.5	27.0	1.1	45.8	0.8	37.7	1.2	42.0	0.4	31.0	1.5	55.8
	RaceHorses	2.0	17.7	1.8	38.5	2.2	33.9	3.1	44.7	1.6	48.4	2.0	45.0
Class D 416×240	BasketballPass	2.2	16.7	1.8	28.9	1.5	33.6	1.7	43.3	1.0	54.6	2.1	56.5
	BQSquare	0.4	22.9	0.4	34.4	0.6	45.1	0.6	35.3	0.2	43.2	1.8	64.1
	Blowingbubbles	2.6	23.4	1.3	33.7	0.7	38.2	1.0	36.1	0.3	41.6	1.8	67.3
	RaceHorses	2.1	12.8	1.8	24.4	1.1	26.6	2.0	36.9	1.1	44.6	2.2	61.9
Class E 1280×720	FourPeople	2.9	56.9	1.4	66.3	1.7	74.1	1.4	69.9	1.1	76.5	1.3	80.4
	Johnny	1.0	57.8	0.9	67.8	1.3	75.7	1.3	73.1	2.3	80.6	1.5	84.5
	KristenAndSara	2.7	51.8	1.3	62.5	1.2	73.1	1.7	71.2	2.0	79.6	1.8	80.6
Average		2.2	32.6	1.2	45.1	1.4	49.6	2.0	51.9	1.2	54.9	2.1	64.2

multiple state-of-the-art methods [12], [14], [18], [21], [28] under both LD and RA test conditions.

V. CONCLUSION

In this paper, a fast and accurate inter mode decision algorithm for the HEVC, called the *unimodal model-based inter mode decision* (UMIMD), is proposed. Based on the observation that a unimodal model exists between the size of different PU modes and their resulted RD costs for each quad-tree partitioned CTU, all the PU modes are classified into several mode classes according to the PU's sizes and these classes are then properly ordered and sequentially checked so that the optimal mode can be early identified by checking when the RD cost starts to arise, indicating that the global minimum of the RD cost was just reached in the previous mode checking. Moreover, to further reduce the computational complexity, an instant SKIP mode termination scheme is developed based on the evaluation of the RD cost of the SKIP mode against a pre-determined threshold. Consequently, the encoding complexity reduction can be further achieved by skipping the computation of the RD cost of the next class when the SKIP mode is the optimal one. Experimental results have shown that the proposed UMIMD algorithm not only significantly reduces the computational complexity but also keeps almost the same coding efficiency, compared with the exhaustive mode decision in the HEVC. Furthermore, it has been demonstrated that the proposed UMIMD algorithm consistently outperforms multiple state-of-the-art methods.

REFERENCES

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

[2] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[3] G. Correa, P. Assuncao, L. Agostini, and L. A. da Silva Cruz, "Performance and computational complexity assessment of high-efficiency video encoders," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1899–1909, Dec. 2012.

[4] B.-G. Kim, "Novel inter-mode decision algorithm based on macroblock (MB) tracking for the p-slice in H.264/AVC video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 2, pp. 273–279, Feb. 2008.

[5] H. Zeng, C. Cai, and K.-K. Ma, "Fast mode decision for H.264/AVC based on macroblock motion activity," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 4, pp. 491–499, Apr. 2009.

[6] H. Q. Zeng, K.-K. Ma, and C. H. Cai, "Hierarchical intra mode decision for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 907–912, Jun. 2010.

[7] E. Martinez-Enriquez, A. Jimenez-Moreno, M. Angel-Pellon, and F. Diaz-de-Maria, "A two-level classification-based approach to inter mode decision in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 11, pp. 1719–1732, Nov. 2011.

[8] J. Y. Lee and H. Park, "A fast mode decision method based on motion cost and intra prediction cost for H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 3, pp. 393–402, Mar. 2012.

[9] X. Shen, L. Yu, and J. Chen, "Fast coding unit size selection for HEVC based on Bayesian decision rule," in *Proc. Picture Coding Symp. (PCS)*, May 2012, pp. 825–828.

[10] T. Zhao, Z. Wang, and S. Kwong, "Flexible mode selection and complexity allocation in High Efficiency Video Coding," *IEEE J. Sel. Topics Signal Process.*, vol. 7, no. 6, pp. 1135–1144, Dec. 2013.

[11] T. Mallikarachchi, D. S. Talagala, H. K. Arachchi, and A. Fernando, "Content-adaptive feature-based CU size prediction for fast low-delay video encoding in HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 3, pp. 693–705, Mar. 2018.

[12] 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, no. 2, pp. 465–470, Feb. 2013.

[13] K.-H. Tai, M.-Y. Hsieh, M.-J. Chen, C.-Y. Chen, and C.-H. Yeh, "A fast HEVC encoding method using depth information of collocated CUs and RD cost characteristics of PU modes," *IEEE Trans. Broadcast.*, vol. 63, no. 4, pp. 680–692, Dec. 2017.

[14] L. Shen, Z. Zhang, and Z. Liu, "Adaptive inter-mode decision for HEVC jointly utilizing inter-level and spatiotemporal correlations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 10, pp. 1709–1722, Oct. 2014.

[15] J. Vanne, M. Viitanen, and T. D. Hamalainen, "Efficient mode decision schemes for HEVC inter prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 9, pp. 1579–1593, Sep. 2014.

[16] K. Goswami, J.-H. Lee, and B.-G. Kim, "Fast algorithm for the high efficiency video coding (HEVC) encoder using texture analysis," *Inf. Sci.*, vols. 364–365, pp. 72–90, Oct. 2016.

[17] 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. Broadcast.*, vol. 60, no. 2, pp. 405–412, Jun. 2014.

[18] 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 Syst. Video Technol.*, vol. 25, no. 3, pp. 422–435, Mar. 2015.

[19] J. Lee, S. Kim, K. Lim, and S. Lee, "A fast CU size decision algorithm for HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 411–421, Mar. 2015.

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

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

[22] K. Duan, P. Liu, K. Jia, and Z. Feng, "An adaptive quad-tree depth range prediction mechanism for HEVC," *IEEE Access*, vol. 6, pp. 54195–54206, Sep. 2018.

[23] S.-H. Jung and H. W. Park, "A fast mode decision method in HEVC using adaptive ordering of modes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 10, pp. 1846–1858, Oct. 2016.

[24] A. Jiménez-Moreno, E. Martínez-Enríquez, and F. Díaz-de-María, "Complexity control based on a fast coding unit decision method in the HEVC video coding standard," *IEEE Trans. Multimedia*, vol. 18, no. 4, pp. 563–575, Apr. 2016.

[25] J. Zhang, S. Kwong, T. Zhao, and Z. Pan, "CTU-level complexity control for high efficiency video coding," *IEEE Trans. Multimedia*, vol. 20, no. 1, pp. 29–44, Jan. 2018.

[26] J. Zhang, S. T. W. Kwong, T. Zhao, and H. H. S. Ip, "Complexity control in HEVC intra coding for industrial video applications," *IEEE Trans. Ind. Informat.*, to be published, doi: 10.1109/TII.2018.2844214.

[27] J. Xiong, H. Li, Q. Wu, and F. Meng, "A fast HEVC inter CU selection method based on pyramid motion divergence," *IEEE Trans. Multimedia*, vol. 16, no. 2, pp. 559–564, Feb. 2014.

[28] Y. Zhang, S. Kwong, X. Wang, H. Yuan, Z. Pan, and L. Xu, "Machine learning-based coding unit depth decisions for flexible complexity allocation in High Efficiency Video Coding," *IEEE Trans. Image Process.*, vol. 24, no. 7, pp. 2225–2238, Jul. 2015.

[29] L. Zhu, Y. Zhang, S. Kwong, X. Wang, and T. Zhao, "Fuzzy SVM-based coding unit decision in HEVC," *IEEE Trans. Broadcast.*, vol. 64, no. 3, pp. 681–694, Sep. 2018.

[30] I.-K. Kim, J. Min, T. Lee, W.-J. Han, and J. H. Park, "Block partitioning structure in the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1697–1706, Dec. 2012.

[31] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G. J. Sullivan, "Rate-constrained coder control and comparison of video coding standards," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 688–703, Jul. 2003.

[32] F. Bossen, *Common HM Test Conditions and Software Reference Configurations*, document JCT-VC-K1100, 11th ITU-T SG16 WP3 Meeting, Shanghai, China, Oct. 2012.

[33] *HEVC Test Model Reference Software (HM)*. Accessed: Oct. 2015. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/

[34] G. Björntegaard, *Calculation of Average PSNR Differences Between RD-Curves*, document VCEG-M33, 13th ITU-T SG16/Q6 VCEG Meeting, Austin, TX, USA, Apr. 2001.

in well-known journals and conferences, including IEEE TIP and TCSVT. His research interests include image processing, video coding, machine learning, and computer vision. He received the Best Paper Award at the Chinese Conference on Signal Processing 2017 and the Best Poster Award at the International Forum of Digital TV and Multimedia Communication 2018. He serves as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE ACCESS, and *IET Electronics Letters*, and as a Guest Editor for the *Journal of Visual Communication and Image Representation*, *Multimedia Tools and Applications*, and the *Journal of Ambient Intelligence and Humanized Computing*. He also serves as a General Co-Chair for the IEEE International Symposium on Intelligent Signal Processing and Communication Systems 2017, as a Technical Program Co-Chair for the Asia–Pacific Signal and Information Processing Association Annual Summit and Conference 2017, as the Area Chair for the IEEE International Conference on Visual Communications and Image Processing, as a Technical Program Committee Member for multiple flagship international conferences, and as a Reviewer for numerous international journals and conferences.

WENJIE XIANG received the B.S. and M.S. degrees in electrical engineering from Huaqiao University, Xiamen, China. She is currently a Lecturer with the Minnan Science and Technology Institute, Fujian Normal University, Quanzhou, China. Her current research interests include image processing and video coding.

JING CHEN (M'17) received the B.S. and M.S. degrees in computer science from Huaqiao University, Xiamen, China, and the Ph.D. degree in control engineering from Xiamen University, Xiamen.

She was a Visiting Scholar with the Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, from 2011 to 2012. She is currently an Associate Professor with the School of Information Science and Engineering, Huaqiao University. Her research interests include image processing and video coding. She is a Traffic Video Technical Committee Member of the China Society of Image and Graphics. She received the Best Paper Award from the Chinese Conference on Signal Processing 2017. She was a Financial Co-Chair and the Session Chair of the IEEE International Symposium on Intelligent Signal Processing and Communication Systems 2017. She serves as a Reviewer for the *Journal of Visual Communication and Image Representation* and *Multimedia Tools and Applications*.

CANHUI CAI (M'06–SM'08) received the B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 1982, the M.S. degree in electronic engineering from Shanghai University, Shanghai, China, in 1985, and the Ph.D. degree in electronic engineering from Tianjin University, Tianjin, China, in 2003.

He was a Visiting Professor with the Delft University of Technology, Delft, The Netherlands, from 1992 to 1993, and with the University of California at Santa Barbara, Santa Barbara, CA, USA, from 1999 to 2000. He has been a Faculty Member with Huaqiao University, Quanzhou, China, since 1984, where he is currently a Professor with the School of Information Science and Engineering. He has authored or co-authored four books and published more than 120 papers in journals and conference proceedings. His research interests include video communications, image and video signal processing, and computer vision.

HUANQIANG ZENG (S'10–M'13–SM'18) received the B.S. and M.S. degrees in electrical engineering from Huaqiao University, Xiamen, China, and the Ph.D. degree from Nanyang Technological University, Singapore.

He was a Postdoctoral Fellow with The Chinese University of Hong Kong, Hong Kong. He is currently a Full Professor with the School of Information Science and Engineering, Huaqiao University. He has published more than 80 papers

ZHANGKAI NI received the M.Eng. degree in communication engineering from the School of Information Science and Engineering, Huaqiao University, Xiamen, China, in 2017. He is currently pursuing the Ph.D. degree with the Department of Computer Science, City University of Hong Kong, Hong Kong. He was a Research Engineer with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from 2017 to 2018. His current research interests include perceptual image processing, deep learning, and computer vision.

KAI-KUANG MA (S'80–M'84–SM'95–F'13) received the B.E. degree in electronic engineering from Chung Yuan Christian University, Zhongli, Taiwan, the M.S. degree in electrical engineering from Duke University, Durham, NC, USA, and the Ph.D. degree in electrical engineering from North Carolina State University, Raleigh, NC, USA.

From 1984 to 1992, he was with IBM Corporation, Kingston, NY, USA, and Research Triangle Park, NC, USA, where he focused on various DSP and very large-scale integration advanced product development. From 1992 to 1995, he was a Technical Staff Member with the Institute of Microelectronics, Singapore, where he is involved in digital video coding and the MPEG standards. He is currently a Full Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. He served as the Singapore MPEG Chairman and the Head of Delegation, from 1997 to 2001. On the MPEG contributions, two fast motion estimation algorithms (Diamond Search and MVFAST) produced from his research group have been adopted by the MPEG-4 standard, as the reference core technology for fast motion estimation. He has published extensively and holds one U.S. patent on fast motion estimation algorithm. His research interests include the areas of digital image/video processing and computer vision, including digital image/video coding and standards, image/video segmentation, denoising, enhancement, interpolation, and super resolution. His research interests on computer vision include image matching and registration, scene analysis and recognition, and human–computer interaction.

He is a member of Sigma Xi and Eta Kappa Nu. He serves as a Technical Program Committee Member, as a Reviewer, and as the Session Chair for multiple IEEE international conferences. He is an Elected Member of three IEEE technical committees, namely, the Image and Multidimensional Signal Processing Committee, the Multimedia Communications Committee, and the Digital Signal Processing Committee. He was the General Chair of organizing a series of international standard meetings (MPEG and JPEG), JPEG2000, and MPEG-7 workshops held in Singapore, in 2001. He is a General Co-Chair of ISPACS2017, ASIPA2017, ACCV2016 Workshop, and VCIP-2013, a Technical Program Co-Chair of ICIP-2004, ISPACS-2007, IIH-MSP-2009, and PSIVT-2010, and the Area Chair of ACCV-2009 and ACCV-2010. He is the Chairman of the IEEE Signal Processing Singapore Chapter, from 2000 to 2002. He has been serving as an Editorial Board Member for several leading international journals in his research area, such as a Senior Area Editor for the IEEE Transactions on Image Processing, from 2016 to 2019, as an Editor for the IEEE Transactions on Communications, from 1997 to 2012, and as an Associate Editor for the IEEE Transactions on Multimedia, from 2002 to 2009, the *International Journal of Image and Graphics*, from 2003 to 2015, the *Journal of Visual Communication and Image Representation*, from 2005 to 2015, the IEEE Transactions on Image Processing, from 2007 to 2010, IEEE Signal Processing Letters, from 2014 to 2016, and the IEEE Transactions on Circuits and Systems for Video Technology, since 2015. He was elected as a Distinguished Lecturer of the IEEE Circuits and Systems Society, from 2008 to 2009.

• • •