Journal of Real-Time Image Processing (2021) 18:603–618
https://doi.org/10.1007/s11554-020-00996-7

ORIGINAL RESEARCH PAPER

Fine-grain complexity control of HEVC intra prediction in battery-powered video codecs

Elahe Hosseini[1] · Farhad Pakdaman[1] · Mahmoud Reza Hashemi[1] · Mohammad Ghanbari[1,2]

Received: 28 December 2019 / Accepted: 27 June 2020 / Published online: 9 July 2020

Abstract
The high-efficiency video coding (HEVC) standard improves the coding efficiency at the cost of a significantly more complex encoding process. This is an issue for a large number of video-capable devices that operate on batteries, with limited and varying processing power. A complexity controller enables an encoder to provide the best possible quality at any power quota. This paper proposes a complexity control method for HEVC intra coding, based on a Pareto-efficient rate–distortion–complexity (R–D–C) analysis. The proposed method limits the intra prediction for each block (as opposed to existing methods which limit the block partitioning), on a frame-level basis. This method consists of three steps, namely rate-complexity modeling, complexity allocation, and configuration selection. In the first step, a rate-complexity model is presented which estimates the encoding complexity according to the compression intensity. Then, according to the estimated complexity and target complexity, a complexity budget is allocated to each frame. Finally, an encoding configuration from a set of Pareto-efficient configurations is selected according to the allocated complexity and the video content, which offers the best compression performance. Experimental results indicate that the proposed method can adjust the complexity from 100 to 50%, with a mean error rate of less than 0.1%. The proposed method outperforms many state-of-the-art approaches, in terms of both control accuracy and compression efficiency. The encoding performance loss in terms of BD-rate varies from 0.06 to 3.69%, on average, for 90–60% computational complexity, respectively. The method can also be used for lower than 50% complexity if need be, with a higher BD-rate.

Keywords Complexity control · HEVC · Intra coding · Pareto optimization

1 Introduction

The latest video coding standard, high-efficiency video coding (HEVC), was finalized by the Joint Collaborative Team on video coding (JCT-VC), in 2013 [1]. Compared to its predecessor H.264/AVC, The HEVC standard reduces the bitrate by half with the same visual quality. The prominent coding performance of HEVC is the result of encoding tools and schemes, such as intra prediction with 35 modes, and the flexible coding structure, coding tree unit (CTU), which consists of coding units (CUs), prediction units (PUs), transform units (TUs), and more prediction modes. However, these coding tools lead to higher power consumption and processing time (about five times higher than H.264/AVC for a certain configuration [2]).

The rapid advances of semiconductor and multimedia technologies have brought many multimedia capable consumer equipment and industrial video applications to use. Many of these applications, such as video conferencing and

✉ Mahmoud Reza Hashemi
rhashemi@ut.ac.ir

Elahe Hosseini
elahehosseini@ut.ac.ir

Farhad Pakdaman
farhad.pakdaman@ut.ac.ir

Mohammad Ghanbari
ghan@ut.ac.ir

1 School of Electrical and Computer Engineering, College of Engineering, University of Tehran, 19466-14886 Tehran, Iran

2 Emeritus Professor at the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, UK

video surveillance, need real-time encoding and are sensitive to delay. Also, portable devices, such as smartphones, and portable computers, have limited processing power and battery capacity. Due to the varying load of computations in different processing elements, shared memory, and battery-dependent power/performance policies at the operating system level of these devices, the available processing power for video encoding varies during the time.

To deal with the video coding complexity, many researchers work on different aspects of lowering encoding complexity which can be categorized into two groups; complexity reduction and complexity control methods. Complexity reduction methods address the problem of reducing the processing time while keeping the encoding performance loss negligible [3–7]. Complexity control approaches, on the other hand, provide flexible and efficient solutions to make trade-offs between rate–distortion (R–D) performance and computational complexity [8–11]. These approaches provide video encoding in any specified target complexity and try to maximize the coding performance within that target complexity. Complexity control is more necessary for real-time and power-constrained video applications, since the complexity reduction methods are usually limited and their performance is highly dependent on the video content.

As an important part of HEVC encoding, intra prediction enables random access to coded frames, prevents the propagation of error, and is a good choice for archiving [12], screen content coding [13], low-delay video coding in complex communication environments and heterogeneous network conditions [8, 14]. As the complexity of intra prediction has specifically increased in HEVC, dealing with its complexity has become very important. Hence, complexity control methods [8], resource allocation schemes in cloud-based systems [15], and encoding time and energy modeling [16] for intra coding have recently become interesting research problems. Moreover, intra coding's contribution to the total encoding complexity of the next-generation video coding standard, versatile video coding (VVC) [17], has increased more than five times, compared to HEVC, which makes dealing with this complexity more pressing.

In this paper, a novel complexity control approach is proposed that for the first time adjusts the complexity via limiting the number of intra modes in each CU (as opposed to the existing approaches that limit CU partitioning). This approach has some advantages over the existing methods: (1) it is more precise and provides a fine-grain control (as the complexity can be adjusted with steps as small as a single intra mode); (2) it is easier to be implemented in the existing video encoders, as it does not require modifying the CTU partitioning algorithm which is the backbone of the HEVC encoder, (3) and it can be employed separately, or on top of existing CU partitioning-based methods for fine-grain control. Several video sequences with various properties have been used to study the effect of encoding parameter selection on the rate–distortion–complexity (R–D–C) space. A set of Pareto-optimal configurations is extracted as the result of this study that corresponds to the best configurations in different complexity quota. A texture-aware fast intra-coding technique is used alongside this optimization to refine the decision process according to the video content. This decision-making technique is integrated into a three-step complexity control. In the first step, a fast online rate-complexity modeling is performed for the sequence under encoding. This model is used to estimate the encoding complexity based on the quantization parameter (QP) and is used to handle the varying bitrate coding. In the second step, a complexity parameter is allocated to each frame, according to the estimated complexity and the target complexity. Finally, the third step assigns a coding configuration to each block of the current frame, based on the allocated complexity, the Pareto analysis, and the frame content.

The main contributions of this paper are summarized as:

A complexity control scheme for HEVC intra prediction is proposed that works based on finding the best set of intra-prediction modes. To the best of our knowledge, this is the first complexity control system that works based on coding modes instead of CU partitioning.

Adjusting the complexity based on the coding modes provides a fine-grain complexity control with lower error rate compared to competing methods. Moreover, it is easier to be implemented on top of existing encoder implementations, as it does not require changing the CTU processing algorithm.

The effect of coding configuration on the R–D–C space of intra coding is explored and modeled, using a Pareto-frontiers analysis. Moreover, a texture-aware fast intra-mode decision [6] is exploited to refine the efficiency of this Pareto-based model. The simplicity of a Pareto-based controller makes it a good choice for both software and hardware encoders.

A simple yet effective encoding complexity-encoding rate model is presented that estimates the encoding time of each video sequence, based on the selected QP. This is useful in handling variable bitrate encoding.

The rest of the paper is organized as follows. Section 2 reviews the related works. Section 3 presents the R–D–C analysis performed. The details of the proposed complexity controller are discussed in Sect. 4. Experimental results verifying the effectiveness of the proposed method are presented in Sects. 5 and 6 concludes the paper.

2 Related works

The problem of the high computational complexity of the HEVC standard has been addressed in several works. Two related tasks can be identified in these works: complexity reduction and complexity control.

2.1 Complexity reduction

Methods in [3–6] try to reduce the number of intra modes that should be assessed to reduce complexity. To this end, the use of dual-tree complex wavelet transform (DT-CWT) and Prewitt operator in finding the dominant edge direction has been studied in [3] and [4], respectively. Hosseini et al. [6] proposed a novel texture analysis method that estimates the intra mode, analyzing the residuals of the planar intra prediction. Moreover, the proposed method can detect the DC/planar modes with the help of the content-adaptive binary arithmetic coding (CABAC) at almost no extra complexity, which leads to further reduction of the rate–distortion optimization (RDO) time. Fast CU decision and fast intra-mode decision based on neighboring blocks are presented in [18] and [19], respectively. convolutional neural networks (CNN) [20, 21] and support vector machines (SVM) [7] are other recent approaches for fast intra-CU partitioning decision. Moreover, hardware-level optimizations for intra coding have been studied in [22–26] where efficient implementations of encoding tools and data reuse techniques are employed to improve energy efficiency.

2.2 Complexity control

Most existing methods [9, 11, 27, 28] control the coding complexity by constraining the maximum CTU depth. Corrêa et al. [27] explored the relationship between the CU depth and coding complexity, and then constrained the maximum largest CU (LCU) depths of certain frames according to those of the previous frames. By adjusting the number of these constrained frames, the target encoding complexity can be met. The complexity-scalable encoder in [28] is capable of adjusting the processing time by limiting the maximum coding tree depth, according to the maximum depth of neighboring and co-located CTUs. In [11], a statistical model is proposed to estimate the coding complexity of each CTU, which helps to restrict the CTU depth range based on the allocated complexity. Jimenez Moreno et al. [9] have proposed a complexity control approach for HEVC which is based on a set of early termination conditions. They obtained thresholds for early termination at different depths via online learning. The methods in [29–31] are based on R–D–C analysis of intensive experiments. A set of configuration pairs has been created by the combination of coding parameters. Each pair has been evaluated on several video sequences to obtain the relationship between R–D performance and complexity. The pairs belonging to the Pareto frontiers are selected to build a lookup table. These complexity control tools can achieve the target complexity ratio by changing the parameter values according to the lookup table. However, ignoring the video content makes this approach prone to loss of efficiency. Grellert et al. [32] propose a complexity metric that measures the complexity according to different used coding tools, such as transforms and sub-pixel interpolations. A proportional, integral, and derivative (PID)-based control system is used to control the coding complexity. A potential issue with CTU-limiting methods is that they require changing the CTU partitioning algorithm and this can be hard in many existing implementations/libraries. As encoding operations are often implemented as recursive algorithms, limiting the operation, especially skipping certain sizes, requires major changes in implementation.

Another approach [10, 33] is to allocate the complexity budget to each CTU according to the visual saliency. Deng et al. [10] present a subjective-driven complexity control (SCC) approach based on the visual attention model. To predict human visual attention, SCC utilizes bottom-up and top-down models to yield the pixel-wise weight map of each video frame, reflecting the saliency values of different pixels. Since the maximum depth of LCU influences the encoding complexity and visual distortion, they formulate a polynomial optimization to minimize visual distortion.

Recently, a few complexity controllers have been proposed that support intra coding. In [34], the method proposed in [27] was extended to adjust the number of constrained frames faster. Zhang et al. [8] explore the relationship between CTU complexity of intra frames and SATD cost, which is shown to be linear. A CTU-level complexity estimation model is proposed according to this relationship. Then the coding complexity is adjusted by selecting a subset of PU sizes for each CTU based on the estimated CTU complexity. For the complexity control accuracy, a feedback-based error elimination scheme is adopted. The subjective-driven approach in [10, 33] also supports intra coding, as well as inter coding. A software defined network (SDN)-based load balancing method for video encoding is presented in [15] that distributes intra coding of videos to several cores and FPGA accelerators. Finally, [16] presents a time and energy modeling for intra coding that can be used to control the intra-coding complexity.

The above-mentioned methods mainly have either or both of these two problems. (1) All discussed methods perform the complexity control by limiting the CU partitioning process and ignore the choice of coding configuration, which can be an accurate means of adjusting the complexity and coding efficiency [6]. This also makes the implementation of these methods harder for the existing encoder designs/libraries, since the CU partitioning is at the core of all HEVC encoder systems. (2) Some existing methods concentrate on the error control and oversimplify the effect of texture characteristics which has a major effect on the choice of encoding. In what follows, a fine-grain complexity controller is proposed that adjusts the complexity using only the encoding configuration and provides a superior coding efficiency by taking into account the video content.

3 R–D–C space exploration

For an efficient complexity controller, an effective understanding of trade-offs between coding complexity and coding efficiency is required. An R–D–C analysis on different choices of HEVC encoder configurations can be very helpful as they have a major impact on both complexity and coding efficiency. Next subsections explain the exploration and modeling of R–D–C space, using the coding configuration of HEVC intra coding.

3.1 Coding configurations for space exploration

To analyze the R–D–C space, first the encoding parameters that have a major effect on both complexity and compression performance are selected, which form the encoding configurations. The rough mode decision (RMD) and RDO processes are the two main steps of HEVC intra coding, and have the largest impact on all-intra-encoder complexity. Therefore, the number of intra modes in RMD and RDO (N_{rmd} and N_{rdo}) are selected for R–D–C analysis. Moreover, these operations affect the performance of different PU sizes differently. While larger PU sizes are less affected by reducing the number of modes, smaller PU sizes are more sensitive to the choice of intra modes.

The baseline HEVC test model (HM) [35] encoder uses 35 modes for RMD in all PU sizes, and {8,8,3,3,3} candidates for RDO in PUs of {4,8,16,32,64} pixels. Since all combinations of these two parameters can create numerous configurations ($\sum_{PU=4}^{64} N_{rmd} \times N_{rdo,PU} \approx 9 \times 10^{10}$ configuration for 5 PU sizes of 64 to 4 pixels), exploration of all these configurations is not practical. As adding or removing a few intra modes to the list of candidates have negligible effect on coding efficiency and coding complexity, a step size of more than one is used to explore the R–D–C space. Experiments show that adding seven modes (step size = 7) to the previous list of candidates in each new configuration provides an acceptable accuracy for a practical complexity control system.

To further reduce the exploration space, without losing much accuracy, some observations from coding several video sequences [6] are notable: (1) since smaller PU sizes (4×4 and 8×8) are more sensitive to the number of intra modes, more modes are checked in RDO of smaller PUs; (2) PUs of 64×64 pixels are very unlikely to be encoded with a diagonal intra mode. Thus, intra modes of RMD for 64×64 PUs are restricted to take three values of {DC/P, DC/P/H/V, all 35 modes} where H and V represent horizontal and vertical modes, respectively; (3) moreover, due to the similar precision in PUs of 4×4 and 8×8 pixels, the same number of candidates are assigned to them. Hence, the number of RMD and RDO candidates for each PU, except for 64×64 pixels, can take values from the following sets.

$$N_{\mathrm{rmd}} \in \{\mathrm{DC/P}, \mathrm{DC/P} + 7\ \text{angular modes}, \mathrm{DC/P} + 14\ \text{angular modes}, \mathrm{DC/P} + 21\ \text{angular modes}, \text{All 35 modes}\}.$$

$$N_{\mathrm{rdo}}\ \text{for PUs of}\{4,8,16,32,64\}\text{pixels} \in \{\{1,1,1,1,1\},\{2,2,1,1,1\},\{3,3,2,2,1\},\{4,4,2,2,2\},\{5,5,2,2,2\},\{6,6,3,3,2\},\{7,7,3,3,2\},\{8,8,3,3,3\}\}.$$

3.2 Model training with Pareto analysis

When the above-mentioned restrictions are applied to parameter selection, 1507 configurations are created for the encoder. In the training process, each configuration is used to encode six high-resolution (1080p) video sequences, including Beauty, Bosphorus, Kimono, Parkscene, RushHour and Sunflower, with QPs of 22, 27, 32 and 37 and the all-intra main configuration [36]. To generalize well on various content, the above-mentioned training videos have been selected such that they cover a wide range of texture and content types. A total of 36,168 encodings are performed accordingly. To reduce the total encoding time required for modeling, 10 frames from the first 100 frames of each video sequence, with strides of 10 frames, were used. Then the Bjontegaard delta rate (BD-rate) and BD-PSNR [37] are calculated to compare each configuration with the baseline HM coding (i.e., the full search). To make complexity of various contents comparable, all measured encoding times are normalized with their baseline HM encoding. In other words, their encoding time is divided by the case of 100% coding complexity.

To analyze and model the R–D–C space, the rate-complexity (R–C) and distortion-complexity (D–C) spaces are explored separately. The blue crosses in Fig. 1a, b show R–C and D–C space for average results of each of the 1507 coding configurations for all the tested video sequences, respectively. It can be observed in Fig. 1 that the configuration 1507 (baseline configuration) has the highest computational complexity and zero degradation; hence, it appears in the rightmost of both plots with complexity of 1 (100%). Similarly, the first configuration which only considers DC/Planar

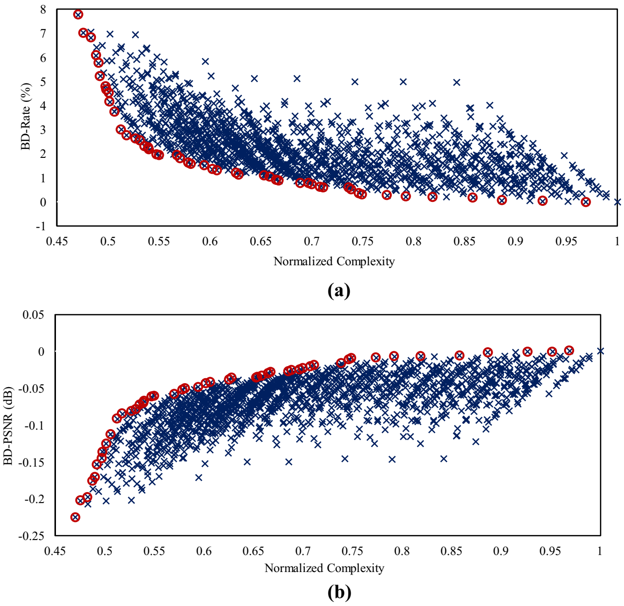

Fig. 1 The average results from the 1507 tested encoding configurations for **a** R–C and **b** D–C spaces. Circles represent Pareto frontier configurations and crosses represent all tested configurations

has the highest degradation and lowest complexity, on the leftmost of both plots.

The R–D–C modeling represents a multi-objective optimization problem where the smallest rate and distortion in each complexity level is desired. This type of problem can effectively be solved by Pareto optimization. To identify the optimal configurations among the 1507 configurations in R–C and D–C spaces, all objectives (rate, distortion, and complexity) are to be minimized at the same time. But due to contradicting objectives, this is not possible in general. To perform a well-balanced trade-off decisions, the configurations which belong to the Pareto frontier in both R–C and D–C spaces are identified.

In case of the R–C space, considering two configuration C_1 and C_2, C_1 dominates (is preferred to) C_2 if each objective (BD-rate and complexity) of C_1 is smaller than the corresponding objective of C_2. Also, in D–C space, C_1 dominates C_2 if it results in a larger BD-PSNR and a smaller computational complexity compared to C_2. The set of configurations that fulfill these conditions belong to the Pareto frontiers and can provide the best average results in terms of R–D–C efficiency within the analyzed configurations. It has been observed that the Pareto analysis might lead to different *optimum values* for each training video; however, the *optimum points* which are the goals of this analysis are almost always the same for all training videos. This is why the average results of all training video sequences have been used for this analysis.

Circles in Fig. 1a, b show the Pareto frontiers in the R–C and D–C spaces. It has been observed that almost all configurations that belong to the Pareto frontiers of one space, are in the Pareto set of the other space as well. This similarity happens due to the nature of the two BD measures used in the analysis, which take into account variations in both the bitrate and PSNR. As illustrated in Fig. 1, the encoding performance naturally diminishes with the decrease of the normalized computational complexity. For complexities above 0.80, degradation of encoding performance is negligible. From 0.80 down to 0.65, the encoding performance diminishes gradually until it reaches a BD-rate increase of 1.04% and a BD-PSNR around 0.033 dB. However, from 0.65 to 0.45, the encoding performance degrades rapidly, resulting in 6.73% increase in BD-rate for 0.2 less complexity.

The reason for this rapid degradation is the sparse intra-mode candidates in lower complexities, which cannot provide an efficient texture modeling. To remedy this, the processing power can be effectively guided to the most probable intra modes, using a fast texture analysis. Section 3.3 employs the method in [6] to improve the quality of modeling.

3.3 Improving the model with texture-aware mode decision

Instead of uniformly selecting angular modes in RMD, the fast mode decision algorithm presented in [6] is exploited here to guide the processing power toward the most probable intra modes. According to [6], the planar mode is proven to be a strong tool to model the plain textures. Therefore, removing the predicted plain textures from the image highlights the dominant high energy texture of the block. To obtain this, first the planar filter is applied to each pixel of the block. Then the residual of planar prediction with respect to the original picture is computed. After obtaining this residual, its energy in the horizontal and vertical directions is measured, to estimate one or two dominant texture directions [6]. Also presents an early termination scheme for DC and planar modes that exploits the context modeling of CABAC. It is shown that when the Hadamard cost of DC and planar are smaller than those of modes with smallest predicted bits based on CABAC's status, DC or planar are most probably selected in the RDO process. This technique has been shown to reduce the complexity with a negligible loss of efficiency.

The content-based decision of this fast mode decision algorithm can improve the encoding performance in low levels of complexity in the proposed complexity control system. Since the DC, planar, horizontal (H) and vertical (V) modes are more frequently chosen for encoding a PU, they have higher priority for lower complexity cases. If the complexity quota allows it, more predicted angular modes (and refinement modes close to them) are added to the list of candidates. Therefore, the following set of values for RMD are designed for the exploration, where 2 angles $\pm i$ denotes the estimated angular modes and i closest modes around them:

$$N_{\mathrm{rmd}} \in \{\mathrm{DC}/P, \mathrm{DC}/P/H/V, \mathrm{DC}/P/H/V + 2 \text{ angles } \pm 1, \mathrm{DC}/P/H/V + 2 \text{ angles} \pm 3, \mathrm{DC}/P/H/V + 2 \text{ angles} \pm 5, \text{All 35 modes}\}.$$

Similar to the previous exploration step, 64×64 PUs are encoded with only DC/P, horizontal and vertical modes. Also, the number of intra modes in RDO is equivalent to the previous R–D–C exploration in 3.1. Accordingly, 1992 configurations are generated in the new exploration phase and overall 47,808 encodings are performed. Repeating the same steps of exploration in 3.2 on all tested configurations from 3.1 and this section (i.e., 1507 + 1992 configurations), the optimum configurations that belong to the Pareto frontier of the entire exploration space are obtained. These optimum points are presented in Fig. 2. Comparing this figure with Fig. 1, it is observed that for higher complexity quota, the same configurations form the first exploration in 3.1 are selected; while for lower complexity quota, the new context-aware configurations are selected. It can also be observed that the degradation in the lower end of plots (leftmost) decreases to almost half, compared with the previous case, which is due to the context-aware decision making.

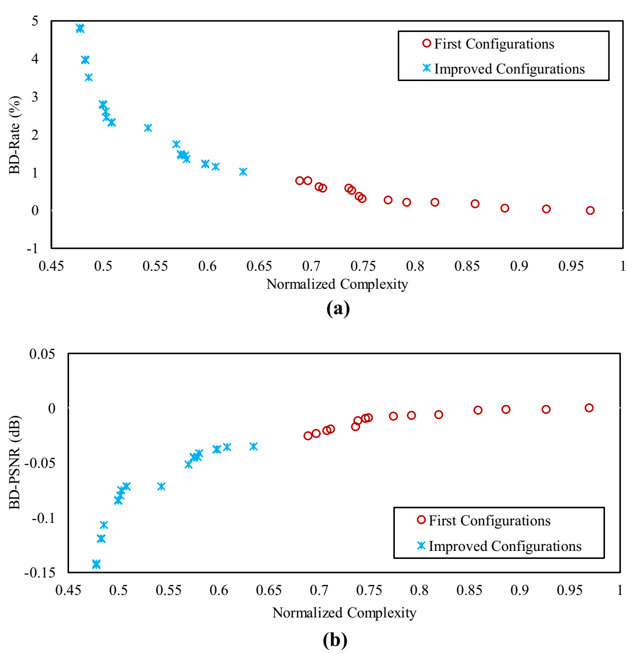

Fig. 2 Final Pareto frontiers among all tested configurations in **a** R–C and **b** D–C spaces

The complexity control algorithm proposed in this work operates by changing the encoder configuration to adjust the computational complexity. To provide a practical control system, 17 final configurations from Fig. 2 with almost equal distances were selected to be used by the complexity control algorithm. Table 1 presents these configurations in descending order of normalized computational complexity. The complexity, BD-PSNR and BD-rate columns present the average results for all training sequences.

Configurations 0–7, shown in white rows of Table 1, adopt the Pareto modes resulting from the fast mode decision algorithm as discussed in Sect. 3.3. Configurations 8–16, on the other hand, shown in gray shadows, select the Pareto modes with uniform distances, as discussed in Sect. 3.2. Table 2 represents the corresponding parameter values for each of these configurations. The RDO and RMD columns indicate the number (or specific modes) of intra modes in RDO and RMD for 4×4 to 64×64 pixels PUs, respectively. $D \pm i$ denotes the estimated direction and i refinement modes around it.

4 The proposed complexity control scheme

This section introduces the components of the proposed control method. The overall framework of this method is summarized in Fig. 3 and includes: (1) rate–complexity modeling, (2) complexity allocation and (3) configuration selection. The rate–complexity modeling component

Table 1 Average results of the selected Pareto configurations

Configuration number	BD-Rate (%)	BD-PSNR (dB)	Normalized complexity (%)
0	4.81	−0.1433	47.77
1	3.97	−0.1189	48.21
2	3.50	−0.1066	48.55
3	2.79	−0.0842	50.01
4	2.16	−0.0671	54.30
5	1.75	−0.0517	57.00
6	1.22	−0.0377	59.82
7	1.02	−0.0315	63.44
8	0.79	−0.0253	68.87
9	0.62	−0.0204	70.78
10	0.30	−0.0091	74.92
11	0.27	−0.0088	77.39
12	0.21	−0.0068	81.89
13	0.18	−0.0058	85.80
14	0.06	−0.0014	88.66
15	0.03	−0.0009	92.65
16	−0.01	0.0005	96.88

Table 2 Selected configurations for complexity control

C	RDO	RMD			
		4×4, 8×8	16×16	32×32	64×64
0	{1, 1, 1, 1, 1}	(DC/P)	(DC/P/H/V)	(DC/P/H/V)	(DC/P)
1	{1, 1, 1, 1, 1}	(DC/P)	10 (D ± 1)	(DC/P)	(DC/P)
2	{1, 1, 1, 1, 1}	(DC/P)	18 (D ± 3)	(DC/P)	(DC/P)
3	{1, 1, 1, 1, 1}	(DC/P)	10 (D ± 1)	10 (D ± 1)	(DC/P)
4	{2, 2, 1, 1, 1}	(DC/P/H/V)	18 (D ± 3)	18 (D ± 3)	(DC/P)
5	{1, 1, 1, 1, 1}	10 (D ± 1)	10 (D ± 1)	10 (D ± 1)	(DC/P/H/V)
6	{1, 1, 1, 1, 1}	18 (D ± 3)	18 (D ± 3)	18 (D ± 3)	(DC/P)
7	{3, 3, 2, 2, 1}	18 (D ± 3)	18 (D ± 3)	18 (D ± 3)	(DC/P/H/V)
8	{3, 3, 2, 2,1}	DC/P + 14	35	35	(DC/P/H/V)
9	{3, 3, 2, 2,1}	DC/P + 21	35	35	(DC/P)
10	{3, 3, 2, 2,1}	35	35	35	(DC/P/H/V)
11	{3, 3, 2, 2, 1}	35	35	35	35
12	{4, 4, 2, 2, 2}	35	35	35	35
13	{5, 5, 2, 2, 2}	35	35	35	35
14	{6, 6, 3, 3, 2}	35	35	35	(DC/P/H/V)
15	{7, 7, 3, 3, 2}	35	35	35	(DC/P/H/V)
16	{8, 8, 3, 3, 3}	35	35	35	(DC/P/H/V)

exploits the relation between QP and encoding time. This is used to improve the accuracy of the proposed complexity controller in the target rate or the scene change. The complexity allocation sets the target complexity for each frame. This module monitors the performance of previous frames and decides the complexity allocation in a way to compensate any unexpected deviations. The last component, configuration selection, chooses the optimum configuration based on a fine-grain content-aware model, to obtain the best possible quality and performance within the allocated complexity. Following subsections detail different parts of this method.

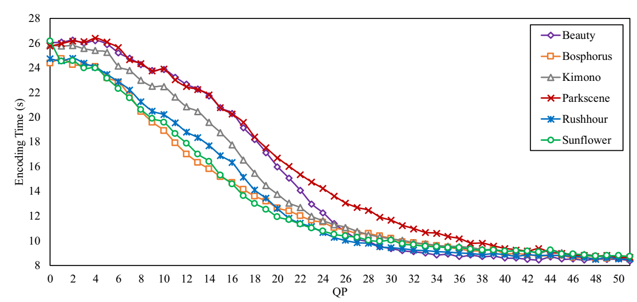

Fig. 4 Average encoding time per frame for different video sequences

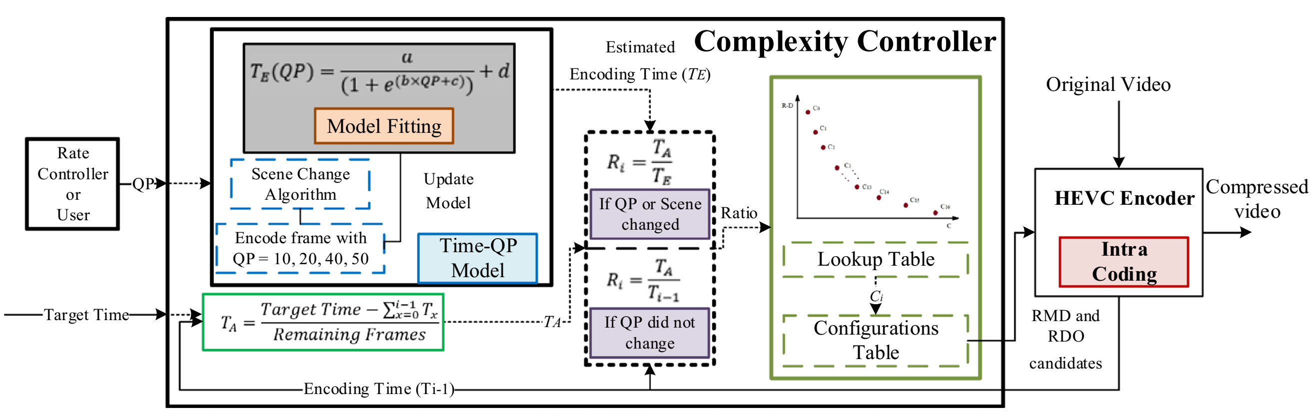

Fig. 3 Overall framework of the proposed complexity control system

4.1 Rate–complexity modeling

Bitrate control is an essential tool for video coding which allows adapting to various network and input conditions. Adjusting QP is one of the main tools to control bitrate. On the other hand, QP has a great impact on the encoding complexity. Thus, to adapt the complexity to the changes of QP, the relationship between QP and encoding complexity is modeled in this paper.

To do so, first all the training video sequences of Sect. 3 were encoded using the HM 16.14 with all QP values (from 0 to 51) and the encoding time has been recorded. All experiments have been performed with the all-intra main HM encoder. The results of these experiments are depicted in Fig. 4, reflecting the average encoding time per frame.

A major conclusion that can be drawn from the experiments is that the complexity of the all-intra coding depends on two factors. (1) QP: controls the number of residual coefficients that are processed by the entropy encoder. (2) Video content: affects the number of bits that are passed to the entropy encoder, especially for mid-range QP values. In the lower values of QP, this dependency is milder, since almost no quantification is made and all coefficients should be entropy encoded. In the higher values of QP, the number of coefficients is largely reduced and just a small number of them are entropy encoded.

It can be observed that all curves present a similar sigmoid-like function, however, with different parameters. Thus, the main function is obtained via model fitting and the exact parameters can be obtained per sequence at the encoding time. Sigmoid function in (1) is used for this, which leads to the highest fitting accuracy. To determine the four parameters of this equation (a, b, c, and d) for a video, which is done at the encoding time, first, four frames are encoded as training frames with QPs 10, 20, 40 and 50 and all-intra configuration. Then, using the four recorded encoding times and QPs, a system of four linear equations is obtained and solved to find the four parameters, which is used to encode all frames. This model needs to be updated at the start of encoding and after a scene change is detected, which is done using the method introduced in [38]. Please note that this step is used (only once) to learn the rate–complexity characteristics of a video scene, and the encoded frames of this step are not sent to the bitstream. Consequently, the QPs are selected to cover a wider range and provide a more accurate estimation.

$$T_{\mathrm{E}}(\mathrm{QP}) = \frac{a}{\left(1 + e^{(b\times\mathrm{QP}+c)}\right)} + d \tag{1}$$

As summarized in Table 3, the coefficient of determination (R^2) [39] values of this model for all the six sequences are above 0.99, which means that almost all variations can be represented by this model and demonstrates the high accuracy of the model.

Table 3 R-square values for different sequences

Sequences	R^2
Beauty	0.9961
Bosphorus	0.9942
Kimono	0.9974
ParkScene	0.9981
Rushhour	0.9973
Sunflower	0.9967

4.2 Complexity allocation

According to the complexity quota and the remaining frames to be encoded in the current task, the controller allocates a specific amount of complexity (equivalent to time in this paper) to the current frame. Assuming that the target complexity is represented as Target Time, the allocated complexity (T_{A}) for the current frame is computed as (2), where the encoding time of previous frames, T_x, is subtracted from the Target Time to determine the available complexity budget.

$$T_{\mathrm{A}} = \frac{\text{Target Time} - \sum_{x=0}^{i-1} T_x}{\text{Remaining Frames}} \tag{2}$$

T_{A} is allocated on a frame by frame basis. This strategy has the advantage that in case the available processing power fluctuates, and thus the encoding time of a frame suddenly changes, this can be considered in the complexity allocation of next frames, which helps finishing the task in the specified time.

The Target Time here is assumed to be an input to our method which is specified from the system's hardware or operating system. Calculating this parameter is out of the scope of our paper, and can be done through power allocation schemes which decide the distribution of system resources between existing tasks or processing modules [40, 41]. In summary, these schemes first estimate the available resources based on information such as hardware status, load of tasks, memory traffic, and expected deadlines. Second, the available power is distributed among existing tasks or processing elements, according to their priorities, real-time requirements, and other system objectives. Finally, the allocated power/resources should be interpreted as a target time for the video encoder, comparing the allocated power to the total requirements, or according to slack times.

Table 4 Lookup table with ratios between normalized complexities of encoding configurations

Next	Current configuration																
	C0	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
C0	1.00	0.99	0.98	0.96	0.88	0.84	0.80	0.75	0.69	0.67	0.64	0.62	0.58	0.56	0.54	0.52	0.49
C1	1.01	1.00	0.99	0.96	0.89	0.85	0.81	0.76	0.70	0.68	0.64	0.62	0.59	0.56	0.54	0.52	0.50
C2	1.02	1.01	1.00	0.97	0.89	0.85	0.81	0.77	0.71	0.69	0.65	0.63	0.59	0.57	0.55	0.52	0.50
C3	1.05	1.04	1.03	1.00	0.92	0.88	0.84	0.79	0.73	0.71	0.67	0.65	0.61	0.58	0.56	0.54	0.52
C4	1.14	1.13	1.12	1.09	1.00	0.95	0.91	0.86	0.79	0.77	0.72	0.70	0.66	0.63	0.61	0.59	0.56
C5	1.19	1.18	1.17	1.14	1.05	1.00	0.95	0.90	0.83	0.81	0.76	0.74	0.70	0.66	0.64	0.62	0.59
C6	1.25	1.24	1.23	1.20	1.10	1.05	1.00	0.94	0.87	0.85	0.80	0.77	0.73	0.70	0.67	0.65	0.62
C7	1.33	1.32	1.31	1.27	1.21	1.15	1.09	1.00	0.92	0.90	0.85	0.82	0.77	0.74	0.72	0.68	0.65
C8	1.44	1.43	1.42	1.38	1.27	1.21	1.15	1.09	1.00	0.97	0.92	0.89	0.84	0.80	0.78	0.74	0.71
C9	1.48	1.47	1.46	1.42	1.30	1.24	1.18	1.12	1.03	1.00	0.94	0.91	0.86	0.82	0.80	0.76	0.73
C10	1.57	1.55	1.54	1.50	1.38	1.31	1.25	1.18	1.09	1.06	1.00	0.97	0.91	0.87	0.85	0.81	0.77
C11	1.62	1.61	1.59	1.55	1.43	1.36	1.29	1.22	1.12	1.09	1.03	1.00	1.95	0.90	0.87	0.84	0.80
C12	1.71	1.70	1.69	1.64	1.51	1.44	1.37	1.29	1.19	1.16	1.09	1.06	1.00	0.95	0.92	0.88	0.85
C13	1.80	1.78	1.77	1.72	1.58	1.51	1.43	1.35	1.25	1.21	1.15	1.11	1.05	1.00	0.97	0.93	0.89
C14	1.86	1.84	1.82	1.77	1.63	1.56	1.48	1.40	1.29	1.25	1.18	1.15	1.08	1.03	1.00	0.96	0.92
C15	1.94	1.92	1.91	1.85	1.71	1.63	1.55	1.46	1.35	1.31	1.24	1.20	1.13	1.08	1.05	1.00	0.96
C16	2.03	2.01	1.99	1.94	1.78	1.70	1.62	1.53	1.41	1.37	1.29	1.25	1.18	1.13	1.9	1.05	1.00

Table 5 Average encoding time error, BD-rate, and BD_PSNR for all test video sequences

Target complexity (%)	BD-rate (%)	BD-PSNR (dB)	Error (%)
50	11.07	−0.56	0.05
60	3.69	−0.19	0.08
70	2.02	−0.11	0.07
80	0.28	−0.01	0.04
90	0.06	0.00	0.06

4.3 Configuration selection

The encoder configuration for the next frame is determined by the ratio (R_i) between the allocated time (T_A) and the encoding time of the previous frame. At the beginning of video coding or after each scene change, where the previous encoding time is not available, the estimated time from the model in (1) is used instead. The complexity controller checks the QP value at the beginning of each frame. If the current QP differs from the one at the beginning of the previous frame, T_E is calculated by (1). This is summarized in (3).

$$R_i = \begin{cases} \frac{T_A}{T_E} & \text{if QP or scene changed} \\ \frac{T_A}{T_{i-1}} & \text{Otherwise} \end{cases} \tag{3}$$

Once R_i is computed, it is used to index a lookup table containing the pre-calculated ratios between the normalized complexities of the 17 selected configurations, to determine the next configuration. The lookup table is marked as Lookup Table in Fig. 4 and is presented in Table 4, where the numbers at the first row indicate the current configuration, numbers in the first column indicate the next frame configuration and the values within each cell indicate the ratio between the normalized complexities of the next and the current configurations. C0 and C16 are the lowest and the highest configurations in terms of computational complexity. For example, if current configuration is C10 and the calculated ratio (R_i) is 0.79, meaning the next configuration should be 21% less complex than C10, the closest value under C10 is 0.8 which means the next configuration based on the table will be C6. As shown in Fig. 4, the new configuration index is used to access a Configuration Table, to find the encoding parameters of Table 2 (number of intra modes in the RMD and RDO processes) which will be used to encode the next frame.

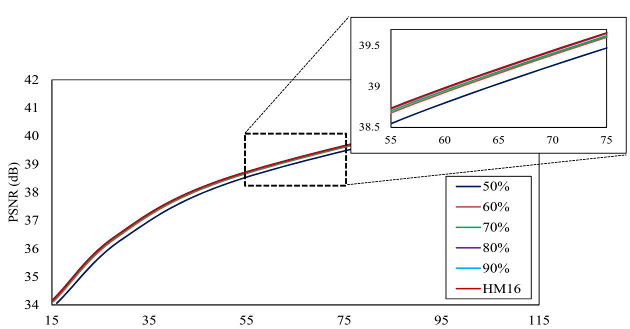

Fig. 5 Rate distortion performance for each target complexity of the Cactus sequence

5 Experimental results

Extensive experiments were conducted to evaluate the proposed method. The proposed method was implemented on top of HM 16.14. The first 100 frames of each video sequence were encoded with QPs 22, 27, 32 and 37, and the all-intra main configuration. The BD-rate, BD-PSNR, and the time error with respect to the target time were measured. The tests have been performed on a computer with a 2.80 GHz CPU and 8 GBs of memory.

5.1 Complexity control accuracy and R–D efficiency results

To evaluate the performance of the proposed method, firstly, the complexity control system was evaluated using five target complexity ratios of 90% to 50%, defined as an encoding complexity ratio between the proposed method and

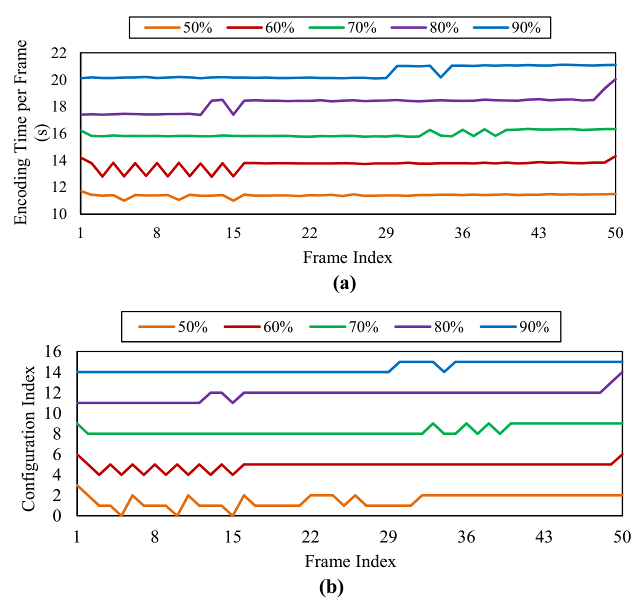

Fig. 6 **a** Encoding time per frame, **b** configuration index for each frame in the PeopleOnStreet and QP = 32

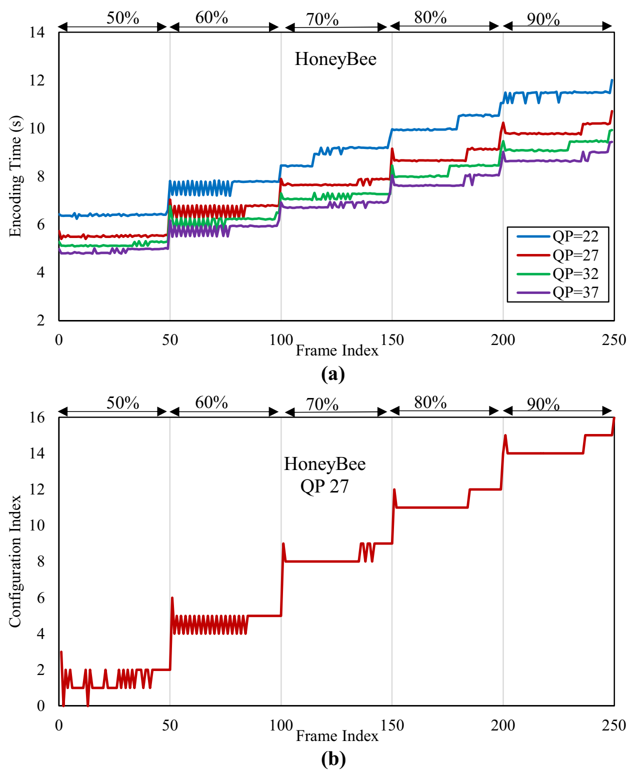

Fig. 7 Dynamic target time scenario for HoneyBee. Target times are given above the figures as a percentage. **a** Encoding time per frame, **b** configurations per frame

baseline HM encoder. In practical implementations of an HEVC encoder, the computational resources available in the encoding platforms will dictate the complexity constraint imposed on the video encoding procedure. To evaluate the proposed complexity controller, 13 sequences (Tables 6, 7) with different resolutions (Class A–E) have been tested. All sequences differ from the training sequences. Table 5 tabulates the average results of these sequences in terms of complexity control accuracy and R–D efficiency. The complexity control accuracy was measured as time error compared with the target complexity ratio, which is calculated by (4):

$$\text{Time Error (\%)} = \frac{T_{\text{Enc}} - \text{Target Time}}{\text{Target Time}} \times 100 \tag{4}$$

Table 5 shows that the encoding time errors vary from 0.04 to 0.08% (averaging 0.06%). Moreover, the compression efficiency results show that BD-rate and BD-PSNR values increase and decrease, respectively, as the target complexity ratio decreases. The loss of compression efficiency for target complexities between 90 and 60% is very small; however, in the lowest complexity ratio, 50%, the loss increases. In this case, the complexity reduction compared to the original encoder is close to the maximum complexity reduction possible using fast mode decision. Thus, naturally the coding efficiency declines. However, this efficiency is acceptable for the low complexity ratio.

Table 5 also reports BD-PSNR which shows the quality degradation compared to HM encoding, in a similar bitrate. The degradations for 90–60% complexity levels are negligible and even for 50% complexity level, it is still very small and with hardly noticeable artifacts. The degradation in all complexity levels are visually insignificant.

Figure 5 demonstrates the R–D performance of the proposed method for different target complexities of the *Cactus* sequence. On the right-hand side, the curve has been zoomed in to show a segment in greater detail. As can be observed, the differences between the target complexities of 60 and 100% are very marginal. When the smallest target complexity, i.e., 50%, is used, the corresponding curve is more visible due to the larger loss of R–D performance, however, still with less than 0.5 dB degradation.

Figure 6 illustrates the operation of the proposed complexity control for the first 50 frames of *PeopleOnStreet* (2600p). Figure 6a shows encoding times per frame, and Fig. 6b shows the evolution of the encoding configuration index, for each of the five different target complexities. The first frame was encoded using the estimated complexity, which shows closer encoding time to the target time; proving the accuracy of the rate–complexity model. In each next frame, the control system adjusts the configuration to deliver the target ratio. Although there are more fluctuations in configurations for lower complexity (50–60%), the encoding times and the used configuration stay almost stable after a few frames.

Moreover, the capability of the proposed method to adapt to dynamic changes of the available processing power was evaluated. To do so, the HoneyBee video sequence was encoded with a scenario where the available processing power (and hence, the target time) changes every 50 frames, from 50 to 90% target time. Hence, the algorithm should adapt the encoding to the new constraint after each change. As depicted in Fig. 7, for all QP values, the algorithm dynamically changes the encoding configurations, so that the target time is met. At the time of changes (vertical lines in Fig. 7a), a small fluctuation of time is observed, which usually resolves within two to three frames. After that, the control algorithm converges to one or two specific configurations and continues coding with those. Some small fluctuations might remain longer in the encoding time, e.g. in 60% target time. In these cases, the target time per frame happens to be between two adjacent configurations (based on Fig. 7b, C4 and C5 in case of 60% target time). Thus, the algorithm sometimes switches between the two configurations to compensate the higher/lower used quota, and remains as close as possible to a zero timing error during the coding process.

Table 6 Experimental results with the target complexity ratio 80%

Sequences	[34]		[10]		[33]		[8]		Proposed	
	BDBR (%)	Error (%)	BDBR (%)	Error (%)	BDBR (%)	Error (%)	BDBR (%)	Error (%)	BDBR (%)	Error (%)
PeopleOnStreet 2560×1600	5.82	−0.3	1.94	1.8	2.08	0.6	0.32	−0.4	0.29	−0.01
Cactus 1920×1080	5.23	−1.15	1.84	2.75	1.84	2.89	0.26	−0.38	0.23	−0.01
BasktballDrive 1920×1080	5.6	−3.59	1.01	2.22	1.36	-0.5	0.29	−1	0.21	−0.04
BasketballDrill 832×480	8.54	−0.68	1.82	2.69	1.93	1.65	0.68	−0.42	0.17	−0.02
BQMall 832×480	8.73	−0.37	2.57	1.12	2.7	1.29	0.91	−0.61	0.33	−0.02
PartyScene 832×480	5.66	1.4	3.55	1.76	3.31	3.11	0.06	0.16	0.41	0
BasketballPass 416×240	7.92	1.96	–	–	–	–	0.63	0.94	0.21	−0.06
BQSquare 416×240	11.63	−1.26	–	–	–	–	0.43	−0.91	0.45	−0.05
RaceHorse 416×240	5.9	2.28	–	–	–	–	0.6	0.42	0.21	−0.11
Johnny 1280×720	9.43	−0.73	0.09	1.43	0.08	2.6	0.53	0.24	0.25	−0.03
KristenAndSara 1280×720	9.96	−0.72	1.16	1.11	1.21	2.35	0.48	−0.24	0.3	−0.04
Average	7.67	1.31	1.75	1.86	1.81	1.87	0.47	0.52	0.28	0.03

Table 7 Experimental results with the target complexity ratio 60%

Sequences	[34]		[10]		[33]		[8]		Proposed	
	BDBR (%)	Error (%)	BDBR (%)	Error (%)	BDBR (%)	Error (%)	BDBR (%)	Error (%)	BDBR (%)	Error (%)
PeopleOnStreet 2560×1600	11.78	−0.47	4.73	1.63	5.03	−0.07	2.35	0.44	3.23	−0.1
Cactus 1920×1080	9.89	−0.09	4.45	1.83	4.44	1.89	2	0.21	1.96	−0.1
BasktballDrive 1920×1080	9.97	−2.73	2.97	1.27	3.33	−0.54	1.38	−0.66	1.97	−0.07
BasketballDrill 832×480	17.67	−1.31	5.5	2.44	6.14	0.79	2.44	0.24	5.37	−0.1
BQMall 832×480	17.41	−0.72	7	0.75	6.98	1.4	3.26	0.03	2.41	−0.13
PartyScene 832×480	12.03	1.35	7.68	0.97	7.43	2.01	1.83	1.31	4.7	−0.05
BasketballPass 416×240	17.3	1.15	–	–	–	–	3.82	0.86	4.13	−0.08
BQSquare 416×240	22.91	−1.44	–	–	–	–	4.63	0.15	6.16	−0.11
RaceHorse 416×240	13.19	3.56	–	–	–	–	3.4	1.19	3.21	−0.09
Johnny 1280×720	19.28	−1.87	2.82	0.93	2.95	2.23	3.04	1.04	4.2	−0.06
KristenAndSara 1280×720	20.33	−1.77	3.6	1.09	3.64	2.6	3.51	0.62	3.24	−0.04
Average	15.61	1.5	4.84	1.36	4.99	1.44	2.88	0.61	3.69	0.08

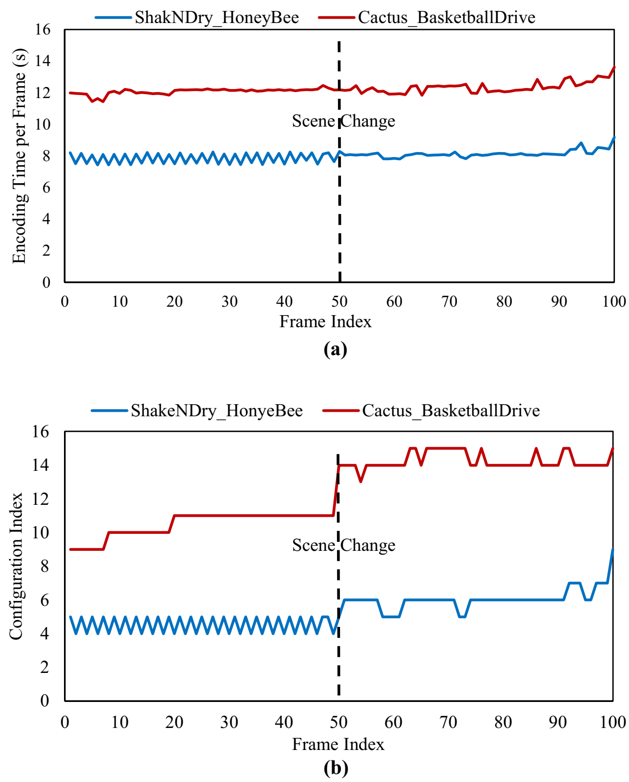

Fig. 8 **a** Encoding times, **b** configurations for each frame for video sequences with scene change and QP = 22

5.2 Comparison with the state-of-the-art methods

In this subsection, the proposed complexity control is compared with four state-of-the-art complexity control methods which support HEVC intra coding [8, 10, 33, 34]. The results of these methods are reported in [8], which uses the same setup for evaluation as used in this paper. The BDBR and time error of these methods for two target ratios, namely 80% and 60%, are provided in Tables 6 and 7. As 240p frames are not reported in [10] and [33], the results of BasketballPass, BQsquare and RaceHorses are left blank for them. The comparison results demonstrate that the proposed method outperforms all other approaches in terms of achieving the target complexity. This is due to the use of intra-encoding modes for complexity control (as opposed to CU partitioning in other methods), which is a finer-grain complexity knob.

It can be observed that the BDBR increases as the target complexity ratio drops for all the five methods. With target complexity ratio of 80% in Table 6, the BDBR and encoding time error values of the proposed method are the smallest (0.28 and 0.03%, respectively), which manifests the superior performance of the proposed method. At the target complexity ratio of 60%, Table 7, the proposed method gains the best time error and the second best BDBR after [8], only marginally. While [8] achieves on average 2.88% BDBR with 0.61% encoding time error (average of absolute values), the proposed method achieves 3.69% BDBR with only 0.08% time error on average (average of absolute values). Moreover, it can be observed in Table 7 that, in higher resolution videos, the two methods achieve similar BDBR, and only in lower resolutions (240p and 480p) [8] gains better BDBR (and worse encoding time error).

Other notable approaches such as [28] and [32] are absent in Tables 6 and 7 as they do not support intra coding, but they can be considered for comparison of timing error. These two works report, respectively, 1.2% and 0.38% timing error in 60% complexity, which confirms the performance of the proposed fine-grain complexity control scheme.

An interesting feature of the proposed method is that it can be employed separately or on top of CU-based methods like [8]. Such a system can use a coarse- to fine-grain adjustment, using [8] and the proposed method, respectively. This benefits the method from high efficiency of [8] at lower complexities and high accuracy of the the proposed method in timing control.

5.3 Complexity control performance with scene change

To evaluate the performance of the proposed method in the presence of scene changes, four video sequences with the same resolutions were concatenated to form two video sequences with scene change. Then the first 100 frames of each composite sequence were encoded. Figure 8a, b presents the encoding times and configurations per frame for these sequences. When scene change is detected at frame 50, the rate–complexity model is updated and the configuration changes according to the new content. The Cactus_BasketballDrive pair is encoded with a target complexity of 80%. It contains a scene change from a very detailed texture and homogenous motion, to a rather simple texture with more intensive motion. The ShakeNDry_HoneyBee sequence is encoded with target complexity of 60% and contains a scene change from a rather smooth texture but with detailed moving objects such as droplets in the air, to a sharper texture with various edge directions, but less details and limited motion. As shown in Fig. 8, although the configuration changes to handle the scene change, the encoding time per frame stays stable for both pairs. This indicates that the proposed complexity control method can smoothly handle scene changes regardless of scene types. The most important factor in scene change is the scene complexity that affects the coding complexity. For Cactus_BasketballDrive, the second scene is much simpler and thus less computationally complex. Hence, as depicted in Fig. 8b, the control algorithm changes the configuration index from 10–11 to 14–15 (i.e., more computations) to keep the previous encoding time. A similar decision with a smaller offset is taken for ShakeNDry_HoneyBee.

6 Conclusion

In this paper, a complexity control method for HEVC intra coding was presented that is useful in power-constraint portable devices and low delay applications. The R–D–C space was explored using encoding configurations and modeled using a Pareto optimization approach, leading to a table of selected coding configurations for each complexity level. A fast mode decision method was used to further enhance the selected configurations. At the controller, first, a complexity estimation model is proposed to estimate the coding complexity of each frame based on the QP (or bitrate). Second, the complexity budget is allocated to each frame proportional to its estimated complexity. Finally, according to the allocated complexity, a coding configuration is determined, which includes a subset of intra modes for RMD and RDO operations. Experimental results have demonstrated that the proposed method accurately controls the encoding complexity, with an average time error from the target complexity of only 0.06%, which is due to employing prediction modes as the complexity knob. The proposed method supports complexity ratios from 100 to 50%. The bitrate increase for 90% to 60% complexity ratio is from 0.06 to 3.69% on average. Also, the quality degradation is negligible and hard to detect visually. While the proposed method provides superior control accuracy and excellent coding efficiency in high to mid ranges of complexity, it can be extended in future alongside (existing or new) CU partitioning-based methods to enhance the performance in lower complexities as well.

References

1. Sullivan, G.J., Ohm, J.-R., Han, W.-J., Wiegand, T.: Overview of the high efficiency video coding (HEVC) standard. IEEE Trans. Circuits Syst. Video Technol. **22**, 1649–1668 (2012). https://doi.org/10.1109/TCSVT.2012.2221191
2. Correa, G., Assuncao, P., Agostini, L., Da Silva Cruz, L.A.: Performance and computational complexity assessment of high-efficiency video encoders. IEEE Trans. Circuits Syst. Video Technol. **22**, 1899–1909 (2012). https://doi.org/10.1109/TCSVT.2012.2223411
3. Pakdaman, F., Hashemi, M.-R., Ghanbari, M.: Fast and efficient intra mode decision for HEVC, based on dual-tree complex wavelet. Multimed. Tools Appl. **76**, 9891–9906 (2017). https://doi.org/10.1007/s11042-016-3584-y
4. Jamali, M., Coulombe, S.: Fast HEVC intra mode decision based on RDO cost prediction. IEEE Trans. Broadcast. (2018). https://doi.org/10.1109/TBC.2018.2847464
5. Liu, X., Liu, Y., Wang, P., Lai, C.F., Chao, H.C.: An adaptive mode decision algorithm based on video texture characteristics for HEVC intra prediction. IEEE Trans. Circuits Syst. Video Technol. **27**, 1737–1748 (2017). https://doi.org/10.1109/TCSVT.2016.2556278
6. Hosseini, E., Pakdaman, F., Hashemi, M.R., Ghanbari, M.: A computationally scalable fast intra coding scheme for HEVC video encoder. Multimed. Tools Appl. **76**, 11607–11630 (2019). https://doi.org/10.1007/s11042-018-6713-y
7. Liu, X., Li, Y., Liu, D., Wang, P., Yang, L.T.: An adaptive CU size decision algorithm for HEVC intra prediction based on complexity classification using machine learning. IEEE Trans. Circuits Syst. Video Technol. **29**, 144–155 (2019). https://doi.org/10.1109/TCSVT.2017.2777903
8. Zhang, J., Kwong, S.T.W., Zhao, T., Ip, H.H.S.: Complexity control in the HEVC intracoding for industrial video applications. IEEE Trans. Ind. Inform. **15**, 1437–1449 (2019). https://doi.org/10.1109/TII.2018.2844214
9. Jimenez-Moreno, A., Martinez-Enriquez, E., Diaz-de-Maria, F.: Complexity control based on a fast coding unit decision method in the HEVC video coding standard. IEEE Trans. Multimed. **18**, 563–575 (2016). https://doi.org/10.1109/TMM.2016.2524995
10. Deng, X., Xu, M., Jiang, L., Sun, X., Wang, Z.: Subjective-driven complexity control approach for HEVC. IEEE Trans. Circuits Syst. Video Technol. **26**, 91–106 (2016). https://doi.org/10.1109/TCSVT.2015.2474075
11. Zhang, J., Kwong, S., Zhao, T., Pan, Z.: CTU-level complexity control for high efficiency video coding. IEEE Trans. Multimed. **20**, 29–44 (2018). https://doi.org/10.1109/TMM.2017.2723238
12. Apple ProRes. https://apple.com/final-cut-pro/docs/Apple_ProRes_White_Paper.pdf
13. Lei, J., Li, D., Pan, Z., Sun, Z., Kwong, S., Hou, C.: Fast intra prediction based on content property analysis for low complexity HEVC-based screen content coding. IEEE Trans. Broadcast. **63**, 48–58 (2017). https://doi.org/10.1109/TBC.2016.2623241
14. Xu, L., Kwong, S., Zhang, Y., Zhao, D.: Low-complexity encoder framework for window-level rate control optimization. IEEE Trans. Ind. Electron. **60**, 1850–1858 (2013). https://doi.org/10.1109/TIE.2012.2190960
15. Sjovall, P., Oinonen, A., Teuho, M., Vanne, J., Hamalainen, T.D.: Dynamic resource allocation for HEVC encoding in FPGA-accelerated SDN cloud. In: 2019 IEEE Nordic Circuits and Systems Conference (NORCAS): NORCHIP and International Symposium of System-on-Chip (SoC), pp. 1–5. IEEE (2019)
16. Rodriguez-Sanchez, R., Alonso, M.T., Martinez, J.L., Mayo, R., Quintana-Orti, E.S.: Time and energy modeling of an intra only HEVC encoder. In: Visual Communications and Image Processing, pp. 1–4. IEEE (2015)
17. Chen, J., Ye, Y., Kim, S.H.: Algorithm description for versatile video coding and test model 6 (VTM 6). Jt. Video Expert. Team ITU-T SG 16 WP 3 ISO/IEC JTC 1/SC 29/WG 11, 15th Meet. Gothenburg, SE (2019)
18. Shen, L., Zhang, Z., An, P.: Fast CU size decision and mode decision algorithm for HEVC intra coding. IEEE Trans. Consum. Electron. **59**, 207–213 (2013). https://doi.org/10.1109/TCE.2013.6490261
19. Marzuki, I., Ma, J., Ahn, Y.-J., Sim, D.: A context-adaptive fast intra coding algorithm of high-efficiency video coding (HEVC). J. Real Time Image Process. **16**, 883–899 (2019). https://doi.org/10.1007/s11554-016-0571-5
20. Zhang, Y., Pan, Z., Li, N., Wang, X., Jiang, G., Kwong, S.: Effective data driven coding unit size decision approaches for

HEVC INTRA coding. IEEE Trans. Circuits Syst. Video Technol. **28**, 3208–3222 (2018). https://doi.org/10.1109/TCSVT.2017.2747659
21. Xu, M., Li, T., Wang, Z., Deng, X., Yang, R., Guan, Z.: Reducing complexity of HEVC: a deep learning approach. IEEE Trans. Image Process. **27**, 5044–5059 (2018). https://doi.org/10.1109/TIP.2018.2847035
22. Ozcan, E., Kalali, E., Adibelli, Y., Hamzaoglu, I.: A computation and energy reduction technique for HEVC intra mode decision. IEEE Trans. Consum. Electron. **60**, 745–753 (2014). https://doi.org/10.1109/TCE.2014.7027351
23. Azgin, H., Kalali, E., Hamzaoglu, I.: A computation and energy reduction technique for HEVC intra prediction. IEEE Trans. Consum. Electron. **63**, 36–43 (2017). https://doi.org/10.1109/TCE.2017.014728
24. Azgin, H., Mert, A.C., Kalali, E., Hamzaoglu, I.: Reconfigurable intra prediction hardware for future video coding. IEEE Trans. Consum. Electron. **63**, 419–425 (2017). https://doi.org/10.1109/TCE.2017.015118
25. Zhang, Y., Lu, C.: High-performance algorithm adaptations and hardware architecture for HEVC intra encoders. IEEE Trans. Circuits Syst. Video Technol. (2019). https://doi.org/10.1109/TCSVT.2019.2913504
26. Kalali, E., Adibelli, Y., Hamzaoglu, I.: A low energy intra prediction hardware for high efficiency video coding. J. Real Time Image Process. **15**, 221–234 (2018). https://doi.org/10.1007/s11554-014-0471-5
27. Correa, G., Assuncao, P., Agostini, L., da Silva Cruz, L.: Complexity control of high efficiency video encoders for power-constrained devices. IEEE Trans. Consum. Electron. **57**, 1866–1874 (2011). https://doi.org/10.1109/TCE.2011.6131165
28. Correa, G., Assuncao, P., Agostini, L., da Silva Cruz, L.A.: Complexity scalability for real-time HEVC encoders. J. Real Time Image Process. **12**, 107–122 (2016). https://doi.org/10.1007/s11554-013-0392-8
29. Correa, G., Assuncao, P., Da Silva Cruz, L.A., Agostini, L.: Encoding time control system for HEVC based on rate-distortion-complexity analysis. In: Proc. IEEE Int. Symp. Circuits Syst. 2015-July, pp. 1114–1117 (2015). https://doi.org/10.1109/ISCAS.2015.7168833
30. Correa, G., Assuncao, P.A., Agostini, L.V., Da Silva Cruz, L.A.: Pareto-based method for high efficiency video coding with limited encoding time. IEEE Trans. Circuits Syst. Video Technol. **26**, 1734–1745 (2016). https://doi.org/10.1109/TCSVT.2015.2469533
31. Kim, J., Kim, J., Kim, G., Kyung, C.-M.: Power–rate–distortion modeling for energy minimization of portable video encoding devices. In: 2011 IEEE 54th International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 1–4. IEEE (2011)
32. Grellert, M., Zatt, B., Shafique, M., Bampi, S., Henkel, J.: Complexity control of HEVC encoders targeting real-time constraints. J. Real Time Image Process. **13**, 5–24 (2017). https://doi.org/10.1007/s11554-016-0602-2
33. Deng, X., Xu, M., Li, C.: Hierarchical complexity control of HEVC for live video encoding. IEEE Access. **4**, 7014–7027 (2016). https://doi.org/10.1109/ACCESS.2016.2612691
34. Correa, G., Assuncao, P., da Silva Cruz, L.A., Agostini, L.: Adaptive coding tree for complexity control of high efficiency video encoders. In: 2012 Picture Coding Symposium, pp. 425–428. IEEE (2012)
35. Rosewarne, C., Bross, B., Naccari, M., Sharman, K.: High efficiency video coding (HEVC) test model 16 encoder description. Document: JCTVC-V1002, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Geneva (2015)
36. Bossen, F.: Common test conditions and software reference configurations. Jt. Collab. Team Video Coding ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, 5th Meet (2011)
37. Bjontegaard, G.: Calculation of average PSNR differences between RD-curves. ITU-T Q.6/SG16, Doc. VCEG-M33, 15th Meet. Austin, Texas (2001)
38. Lee, J., Shin, I.H., Park, H.W.: Adaptive intra-frame assignment and bit-rate estimation for variable GOP length in H.264. IEEE Trans. Circuits Syst. Video Technol. **16**, 1271–1279 (2006). https://doi.org/10.1109/TCSVT.2006.881856
39. Glantz, S.A., Slinker, B.K., Neilands, T.B.: Primer of Applied Regression and Analysis of Variance. McGraw-Hill, New York (1990)
40. Zoni, D., Cremona, L., Fornaciari, W.: All-digital control-theoretic scheme to optimize energy budget and allocation in multicores. IEEE Trans. Comput. **69**, 706–721 (2020). https://doi.org/10.1109/TC.2019.2963859
41. Khan, M.U.K., Shafique, M., Henkel, J.: Power-efficient workload balancing for video applications. IEEE Trans. Very Large Scale Integr. Syst. **24**, 2089–2102 (2016). https://doi.org/10.1109/TVLSI.2015.2504415

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Elahe Hosseini received her B.Sc. and M.Sc. degrees both in computer engineering from Amirkabir University of Technology, and University of Tehran, Tehran, Iran, respectively. Her research interest includes power-aware and power-efficient video encoding systems.

Farhad Pakdaman received his B.Sc. and M.Sc. degrees both in computer engineering from the University of Mazandaran and the University of Tehran in 2011 and 2013, respectively. He received his Ph.D. in computer engineering from the University of Tehran in 2019. His research interests include multimedia systems and applications, video processing and compression, high performance computing, power-aware computing, and application-specific architectures. From 2014 to 2017, he has served as a lecturer at the University of Mazandaran. From January 2018, he has also been a researcher at the Multimedia Research Group (MRG), Tampere University, Finland. He has been a reviewer of IEEE TCSVT, IEEE Access, and Multimedia Tools and Applications.

Mahmoud Reza Hashemi received his B.Sc. and M.Sc. degrees in electrical engineering from the University of Tehran, and pursued his Ph.D. degree at the University of Ottawa. He is currently a tenured Associate Professor at the School of Electrical and Computer Engineering of the University of Tehran. He is also the co-founder and Director of the Multimedia Processing Laboratory (MPL). His research interests include multimedia systems and networking, computer architecture, and recently, cloud gaming. He is an Associate Editor of IEEE Transactions on Circuits and Systems for Video Technology. Dr. Hashemi is a senior member of the IEEE and a recipient of the ICME Quality Reviewer Awards in 2011 and 2013.

Mohammad Ghanbari is a professor of video networking at the University of Tehran, Iran and Emeritus Professor at the university of Essex, UK. He is best known for his pioneering work on two-layer video coding for ATM networks (which earned him an IEEE Fellowship in 2001), now known as SNR scalability in the standard video codecs. He has registered for 13 international patents on various aspects of video networking and was the co-recipient of A.H. Reeves prize for the best paper published in the 1995 Proceedings of IET on the theme of digital coding. He is the author of seven books and his book, An Introduction to Standard Codecs, received the year 2000 best book award by IET. Prof. Ghanbari has authored or co-authored more than 700 journal and conference papers, many of which have had a fundamental influence in the field of video networking. He was one of the founding associate editors to IEEE Transactions on Multimedia (IEEE-T-MM) and had served from 1998 to 2002. On January 2014, he was honored as Life Fellow of IEEE.