



A computationally scalable fast intra coding scheme for HEVC video encoder

Elahe Hosseini¹ · Farhad Pakdaman¹  · Mahmoud Reza Hashemi¹ · Mohammad Ghanbari^{1,2}

Received: 15 April 2018 / Accepted: 19 September 2018 /
Published online: 29 September 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

The High Efficiency Video Coding (HEVC) adopts 35 intra prediction modes to provide a more precise intra prediction. As a result, HEVC intra coding is much more complex. To provide a flexible solution for battery-powered devices, which include varying power budgets based on battery level, this paper jointly employs statistical distribution of different intra modes, and texture analysis. While the statistical information specifies the most efficient modes for extremely low power budgets, for moderate to high power budgets the best angular modes are predicted by analyzing the residual of the Planar mode, which is demonstrated to represent the dominant edges of the block. Moreover, a scheme is presented that further reduces the complexity by predicting the more important DC and Planar modes using the information provided by the entropy coder. The proposed method imposes almost no computational overhead since it utilizes the existing internal operators of HEVC. Ultimately, an intra prediction scheme with seven predefined levels of complexity is presented that provides coding for various power budgets and devices with various capabilities. The highest level of this scheme provides ~23% total encoding time reduction compared to HM, in all-intra configuration, with only 0.78% penalty on BD-Rate for high end devices. While the lowest level makes encoding possible for extremely low power devices and low battery plans with ~52% time reduction (almost the entire share of intra coding) at the cost of 9.6% BD-Rate, which is shown to be nearly minimal for this level of processing power.

Keywords HEVC · Intra prediction · Power-aware video coding · Video compression

✉ Mahmoud Reza Hashemi
rhashemi@ut.ac.ir

¹ School of Electrical and Computer Engineering, University of Tehran, Tehran 14399571, Iran

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

1 Introduction

The rapid growth of High Definition (HD) and Ultra HD (UHD) content imposes a great challenge for video encoding. Efficient transmission of these contents requires a high compression ratio such that the state-of-the-art H.264/AVC [32] standard fails to provide. To manage the high bitrate, the Joint Collaborative Team on Video Coding (JCT-VC) introduced the High Efficiency Video Coding (HEVC) [27]. The HEVC standard offers 50% reduction in bit rate with equal quality compared to H.264/AVC. With 35 luma intra prediction modes (33 directional modes plus DC and Planar), HEVC intra prediction is much more complex than H.264/AVC, with only nine modes. Also, instead of Macro Block (MB) the HEVC standard adopts Coding Tree Unit (CTU), which can recursively split the picture blocks, leading to Coding Units (CUs) of 64×64 to 8×8 pixels. These capabilities considerably improve the intra encoding efficiency and provide 23% bit-rate reduction on average compared to that of H.264/AVC [31]. To achieve optimal coding efficiency, the intra prediction unit needs to test all these 35 prediction modes for different CU sizes, demanding high computational complexity. This consequently leads to more power consumption and processing time. Complexity analysis of HEVC Test Model (HM) [19] indicates it can be up to 502% more complex compared to H.264/AVC [5].

Many researchers have been working on different methods to reduce the computational complexity of the HEVC encoder. Most of these works have been addressing the problem of reducing the processing time which is a critical aspect in platforms that entail real-time encoding of video. However, multimedia capable devices have variable power resources, from low-price smart phones with software-based video encoders, to high end smart phones with hardware cores for video encoding, to powerful broadcast equipment with dedicated professional encoder boards. Moreover, in battery-powered devices, the power allocation policy changes based on the battery level; when fully charged, the device might work in full power to deliver best performance, however in low power the device works with lower frequency to extend battery time. Therefore, a single solution cannot be efficient for all devices and all power budgets (The word power denotes both processing power and battery power in the rest of this paper). Thus, the bigger challenge is to provide a computationally scalable method that can cover a wide range of devices. Such a method can provide different configurations/plans for different levels of power, such that even low processing power devices which are not able to perform HEVC encoding, can encode videos with acceptable quality degradations. Or battery powered devices can continue video capturing for longer time. Also, more powerful devices will have the chance to trade power for better coding efficiency.

The above problem has been addressed in this paper by jointly investigating the statistical occurrence of intra modes and texture analysis, so that low power and moderate power options can be covered as well as high performance power policy. The statistical information provides a mode selection scheme which checks the most frequent intra modes first, to maximize cost-efficiency for low power devices. Also, the strong ability of the planar mode in modeling the plain texture is exploited. Its residual, which represents noise-free dominant texture, is analyzed to estimate the dominant texture direction of the block, leading to the best angular modes for low to moderate power budgets. With the knowledge that the best mode is near this direction, different number of refinement modes are checked around it based on the available power, leading to a computationally scalable intra coding scheme for HEVC. Based on this scheme, seven pre-defined configuration modes can be selected for various power budgets. Furthermore, a scheme is proposed that can estimate the selection of DC and Planar modes,

which are the most frequent modes, to further decrease the computational complexity of the Rate-Distortion Optimization (RDO). To do so, Sum of Absolute (Hadamard) Transformed Difference (SATD) of DC and Planar, are compared with SATD of the angular modes with maximum chance of being selected. These modes are distinguished using the internal mode-bit estimator of the Content-Adaptive Binary Arithmetic Coding (CABAC), which predicts the required coding bits for each mode in advance, based on information from neighboring Prediction Units (PUs). All the operations in the proposed method exploit the built-in HEVC modules, which are inevitable parts of the encoder, so that the overhead is almost zero. The main contributions of this paper are:

- A method is proposed to early estimate the selection of DC and Planar modes, prior to RDO, by comparing their estimated CABAC bit-modes against the potentially best angular modes. When DC or Planar are inferred as the best modes, the expensive RDO will be turned off to save processing power. This action can save up to 23% of the encoding time, with a penalty of only 0.74% increase in term of Bjontegaard Delta Rate (BD-Rate) [1].
- It is demonstrated that the residuals of the Planar mode can effectively represent the dominant high energy texture. Based on this, a novel texture analysis method is presented that estimates the best angular modes for each block.
- The proposed method utilizes the built-in HEVC functions, like the filters associated with the Planar mode, and the estimated bit-modes, based on built-in tables of the CABAC. Therefore, almost no computational overhead is imposed. Moreover, in hardware-implementations, this feature leads to more efficient resource utilization and avoids area overhead.
- Utilizing the flexible nature of the proposed method and statistical occurrence of the intra modes, a computationally scalable method with seven pre-defined complexity levels is presented, which is applicable to devices with various levels of processing power. This includes total encoding time reduction of ~23% to ~52% compared to HM (Since the method has been focused on intra coding, all experiments have been carried out with all-intra configurations [2]). Since all the seven configurations are based on the same method, a device can easily switch between these configurations without any need for modifications in hardware/software. This makes the proposed method suitable for dynamically adjusting the processing power in battery-powered environments.

The rest of the paper is organized as follows. Section 2 presents an overview of intra mode decision in HEVC. A review of related works is carried out in Section 3. Section 4 introduces our fast intra mode decision algorithm and presents a computationally scalable method. Section 5 is dedicated to the experimental results and Section 6 concludes the paper with some remarks.

2 Intra mode decision in HEVC

Intra unit exploits spatial correlation within each PU and its neighboring pixels for better prediction. For efficient removal of spatial redundancy to achieve higher compression, some new features are defined in HEVC such as different structures (CTU, CU, PU and TU). HEVC uses quadtree and recursive structure to split CUs. Each CU can be divided into four PUs and intra prediction is performed for each PU. The size of a CU can be from 64×64 to 8×8 pixels,

therefore the size of PUs can range from 64×64 to 4×4 pixels. Accordingly, HEVC performs intra prediction for blocks of 64×64 to 4×4 pixels. The Rate Distortion (RD) cost is calculated for all possible block sizes to find out an optimal block partitioning.

As mentioned earlier, the HEVC encoder supports 35 modes, where 33 of them are angular and two others are the DC and Planar modes. There are two options available in HM encoder to select the best mode: Fine Mode Decision (FMD), and Rough Mode Decision (RMD) followed by a selective Rate Distortion Optimization (RDO). In FMD, RDO should be carried out over all the 35 modes. Considering that all block sizes from 64×64 to 4×4 pixels need to be tested, the RDO cost is calculated $35 \times \left(\sum_{i=0}^4 4^i \right) = 11935$ times for each FMD, which is significantly high and is not practical.

On the contrary, RMD is a fast mode decision algorithm which is followed by a *selective* RDO process. Figure 1 shows this process in the HM encoder. First, the Hadamard transform, which has a relatively low computational complexity, is used instead of the Discrete Cosine Transform (DCT) to check all the 35 modes, and N best modes based on minimum SATD cost are selected. Then the RDO process is applied only to the N selected modes, which requires much less computations. N is predefined for each PU size and is set to $\{8, 8, 3, 3, 3\}$ for blocks of 4×4 , 8×8 , 16×16 , 32×32 and 64×64 pixels, respectively. The number of candidates for smaller blocks is larger, because these blocks are more prone to noise and thus deciding the best mode requires a finer level of analysis. Consequently, the RDO is calculated $(1 \times 3 + 4 \times 3 + 16 \times 3 + 64 \times 8 + 256 \times 8) = 2623$ times for each block in RMD, which is much smaller compared to FMD. Equation (1) presents the cost function used for RMD based on the SATD.

$$cost = SATD + \lambda_{pred} \times B_{pred} \quad (1)$$

In Eq. (1), λ_{pred} is a Lagrange factor and its value depends on the intra mode. B_{pred} is the coding bits representing the number of transmission bits for prediction mode. CABAC calculates the coding bits and updates these values after each prediction. In calculating SATD, first, residual of prediction for each pixel of the current PU is calculated using (2) where $Orig(x,y)$ and $Pred(x,y)$ represent original and predicted values per coding unit, respectively. Then SATD is calculated from Eq. (3).

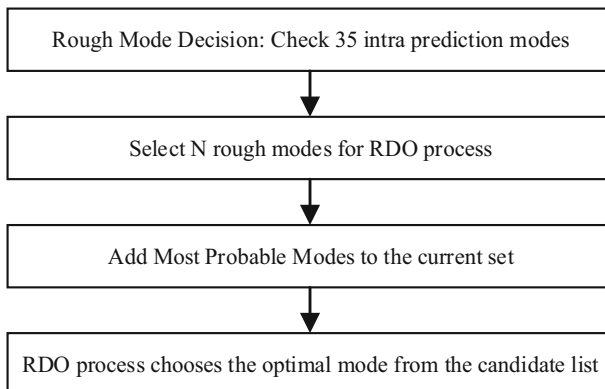


Fig. 1 The process of intra prediction in HM

$$Diff(x, y) = Orig(x, y) - Pred(x, y) \quad (2)$$

$$SATD = \sum_{x,y} |HAD(Diff(x, y))| \quad (3)$$

The next step after RMD is adding the Most Probable Modes (MPM) to the list of candidates, which are predefined modes including the optimum mode of the current PU's left and up neighbors. RD cost function (J_{mode}) is used for the final mode selection through Eq. (4).

$$J_{mode} = (SSE_{luma} + W_{chroma} \times SSE_{chroma}) + \lambda_{mode} \times B_{mode} \quad (4)$$

Where SSE (Sums of Squared Errors) is the sum of squared difference between the current CU and the matching blocks. B_{mode} represents the number of bits including intra prediction mode, splitting information and residuals.

Our investigations reveal that in all-intra configuration, about 15% of the total time is spent in the RMD stage of HM. This is due to checking a large number of intra modes. Also, RDO contributes to about 40% of the encoding time. According to these findings, it can be inferred that using fast mode decision schemes, a maximum of nearly 55%-time reduction can be achieved. Clearly, such a method would lead to a lower coding efficiency due to lack of enough intra analysis. A better trade-off would be to reduce a part of this complexity for better acceptable coding efficiency. The proposed method in this paper presents a scalable method that can provide acceptable coding efficiency for various devices with different power budgets, including time savings between nearly 20 and 50%.

3 Related works

Several fast intra prediction algorithms have already been proposed to reduce the computational complexity of intra coding units, and most of them have reduced the encoding time. These works can be classified into three main categories: fast intra mode decision [3, 4, 7, 8, 10, 11, 13, 17, 18, 22, 24, 25, 30, 33, 34, 36, 37, 40, 41], fast CU size decision [9, 14, 20] and mixed methods of jointly considering intra mode and CU size to reduce the overall complexity [12, 15, 23, 26, 35, 38, 39]. Some proposed works have utilized preprocess functions to reduce the number of intra modes. Laplacian and Sobel are well known filters which have been applied to PUs to analyze the texture and find the prediction mode with edge detection algorithms. Jiang et al. [11] calculate the amplitude and the angle of the gradient for current PUs' pixels and use the gradient-mode histogram to select the most probable modes and reduce the set of candidate prediction directions for RMD process. Also gradient histogram has been used in [3] to reduce prediction modes. Methods in [4, 10, 25] have reduced the number of tested modes in RMD using the Sobel operator. However, Jamali et al. [10] employed an improved edge detector based on Sobel operator and binary classification to reduce the number of candidates of the RD cost calculation. The proposed method in [25] employs the Sobel edge operator to determine the candidate intra mode of 3D-HEVC. In our previous work [17], we have analyzed the energy distribution in sub-bands of a dual-tree complex wavelet Transform (CWT) to model the edges and estimate the dominant edge direction for each PU. It was demonstrated that due to the high reliability of this predicted direction, checking only three modes around this mode, plus DC and Planar, are enough to maintain the intra frame compression efficiency of HEVC but in a much faster way.

Some works [7, 13, 24, 30, 33, 34, 37, 40, 41] have been conducted to reduce the computational complexity of choosing the best intra mode exploiting statistical information. Zhu et al. [40] propose a fast algorithm for intra prediction mode with three different cases using a decision tree. The decision tree is generated by the variances of the current PU and those of its left and above neighbors. In case one, when variance values are smaller than a threshold and PU is very smooth, thus RMD checks only DC and Planar modes. In case two, there are 19 intra modes available in the RMD stage. If none of those cases are satisfied, all the 35 intra modes are tested. In addition, the number of modes in the RDO process has been reduced. An AR (Auto Regression) model has been proposed in [13] based on the relationship between RMD and FMD to reduce the number of candidates for the FMD stage. The distribution of the SATD costs and adjusting parameter, control the number of candidate modes for FMD. This parameter is selected according to the upper and left blocks around the current block. Zhu et al. [41] used both the sum of absolute difference (SAD) values and the SATD values to determine the modes to be tested in the RDO step. In their proposed fast intra mode decision, firstly SAD costs of all 35 modes were calculated and those modes which met a threshold criterion have been selected for the next step. Then, SATD costs of the reserved prediction modes were calculated and if it was smaller than a second threshold, the mode was chosen as a candidate mode of RDO. The two threshold values have been defined based on the statistical analysis of a large number of video sequences. The works in [34] and [24] applied five different filters to detect the dominant edge. Authors in [24] categorized intra modes into five types (0° , 45° , 90° , 135° and non-directional); but [34] improved the mode classification into four types (0° , 45° , 90° , 135°) based on dominant edge assent (DEA) and checked non-directional modes in each set. The dominant edge was then decided by the difference and standard deviation of DEA and a subset of prediction modes (seven or eleven). These were selected as the candidate modes for the RMD process. Yan et al. [33] reduced computational complexity by early termination when the cost of one mode was smaller than a threshold. Then, they merged the adjacent modes according to their angles into several groups. When there were two or three groups, a pixel-based edge detection algorithm was applied to choose the one with the least edge absolute difference. But if all angular modes merged into one group, the mode with the minimum RMD cost and its two neighbors were selected as candidate modes.

Regarding the fast intra mode decision [8, 18], propose to reduce the number of tested modes by sparse search. Fini et al. [18] introduced a two stage algorithm. The even numbered directional modes are checked in the first stage, along with Planar and DC. The costs of odd numbered modes are estimated by the average of the costs of their two even neighboring modes. Therefore, the complexity of RMD testing modes is reduced from 35 to 19 modes. In the second stage, the number of intra modes in the RDO stage is reduced accordingly. Authors in [8] proposed fast intra-algorithms based on the refinement of general directions and the correlation between neighboring blocks. They firstly, calculate the SATD cost of 9 general directional modes (2, 6, 10, 14, 18, 22, 26, 30 and 34) and then an optimal adjacent mode (OAM) list is built according to the positions of minimum and second minimum SATD cost of directional modes in the subset.

Zhang et al. [36] developed several algorithms to reduce complexity. The Hadamard transform on 2:1 down-sampled prediction residual was applied to derive SATD for RMD. They then reduced rate-distortion optimized quantization complexity based on SATD cost and mode distances.

Several fast algorithms are based on determining the coding tree structure and CU size. In [9], the spatial correlation of pixels in the current block is used to determine the optimal CU size. When the mean and variance of pixels are less than thresholds, which were determined experimentally, the splitting of the corresponding block is seized. Ruiz et al. [20] developed a partitioning algorithm for low complexity intra coding based on a decision tree. The decision tree was trained off-line using Data Mining models. In [14], texture complexity is used for partitioning decision. First the Sobel edge detector was used to measure the texture complexity and then thresholding techniques were used to terminate further CU splitting whenever possible.

Other approaches optimize both the intra mode decision and the CU partitioning. Song et al. [26] employ an adaptive early CU depth decision based on an adaptively selected threshold which is obtained by analyzing the first frame in the group of pictures (GOP). They apply an orientation gradient method to analyze PUs and reduce the number of candidate modes in RMD. In [38], a fast algorithm for intra coding is introduced in two micro and macro levels. In the micro level, it tries to reduce the encoding complexity of the RMD stage of the standard encoder using a Progressive Rough Mode Search (pRMS) method. pRMS restricts the 35 intra modes to only 11 modes, then RMD calculates the SATD costs to fill the intra mode candidates list. Finally, the algorithm computes RDO costs for the best three candidate modes. In the macro level, time reduction is gained by an early CU size decision, using required bits for intra prediction modes coding (IPMC), which is provided by the CABAC engine. Yao et al. [35] also exploit the distinct difference in the number of IPMC bits between the MPM and non-MPM modes in their fast intra mode decision method, by reducing the RDO computation. When IPMC bits of MPM modes are less than those of non-MPM modes, more probable MPM modes will be chosen as the optimal prediction mode. Additionally, the proposed method estimates the cost of further CU partitioning with the value of average IPMC bits to reduce the complexity of CU depth decision.

A few power aware solutions have also been presented to control the complexity and distortion of the encoder by adjusting the encoding parameters. In [21], the parameters which have a more significant impact on both complexity and quality are extracted. Then, a set of encoder configurations are defined for each complexity level. Furthermore, the paper proposes a Distortion-Complexity model. Correa et al. [6] however, performed R-D-C analysis using a set of configurations which include encoder parameters and early termination schemes. Then, to adjust the computational complexity, the best configurations are selected through Pareto models. The method has also employed a medium-granularity time control (MGTC) system and fine-granularity encoding time control (FGTC) to keep encoding time per GOP below a target value and guarantee high accuracy.

Although the above research solutions reduce the computational complexity of the HEVC intra encoder, but they are each suitable for a narrow range of computational/power budget and cannot perform outside of that range. The mentioned power-aware methods [6, 21] can provide proper configurations for each power level. However, they are mainly designed for the motion estimation operations and do not consider the intra coding operations. Furthermore, since these methods consider only encoder-provided options (as opposed to complexity constrained methods in this paper), their RD loss can be significant for low power devices. In the following sections, we propose a method, which can detect the DC/planar modes and angular directions of each PU, using the internal information of the encoder. This information is then used to provide a scalable power-adaptive scheme with seven different levels of complexity for intra coding.

4 Computationally scalable intra coding scheme

In recent years, video encoding has become a built-in feature of almost all battery-powered devices. Most of these devices can control and change their power allocation and consumption, according to the remaining battery level. Similarly, consumers have learned to set their quality expectation according to battery level (quality vs. encoding duration) to extend battery time. This might include changing the recording quality, or screen brightness to a moderate level, when the battery is draining. As a result, a computationally scalable scheme, that can provide a suitable solution based on the available power budget, is vital.

The main challenge for a computationally scalable scheme is to perform well under low battery to middle range power budgets, since normally devices have no difficulty performing in full power, where enough processing power is available to perform the necessary operations. Under extremely low power budgets when only a few intra modes can be checked, it is critical to select the most important intra modes to deliver the best possible efficiency and quality. Our investigations show that random selection of few modes in this situation can lead to more than 20% increase in BD-Rate, while selecting the proper modes based on statistics, can reduce this number to half. Furthermore, in moderate power budgets, it is helpful to have an estimate of the best intra direction, such that refining this direction would converge to the best mode much faster. Such that, the intensity of refinement around this direction can be decided based on the available power. Accordingly, this paper jointly employs statistical analysis of intra modes and texture analysis.

Figure 2 shows an overall view of the scheme presented in this paper. While statistical analysis determines the most efficient modes for low power, texture analysis provides the main

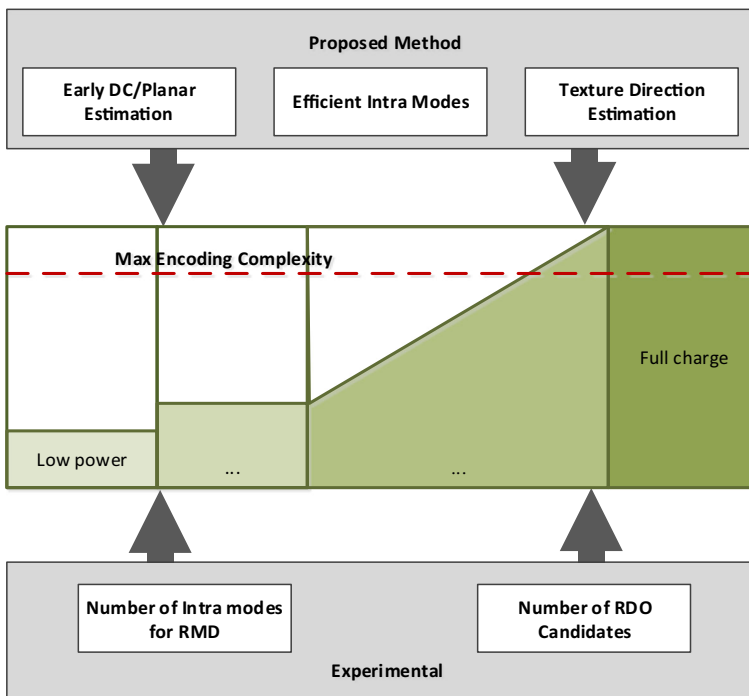


Fig. 2 An overall view of the proposed intra coding scheme

direction, around which search continues in case more power is available. Different configurations are presented to set the number of modes for RMD, and accordingly the number of RDO candidates, based on predefined power limitations. These configurations have been determined experimentally to balance between coding efficiency and computational complexity. Furthermore, as depicted in the figure, the early estimation of DC and Planar modes provides the possibility to turn off the RDO, which decreases the overall encoder complexity in almost all power configurations.

For experimental analysis, extensive sets of experiments have been performed on 11 video sequences, listed in Table 8 below. For this purpose, the video sequences were encoded with Quantization Parameters (QP) equal to 22, 27, 32 and 37 to measure BD-Rate and BD-PSNR. Considering many different combinations for number of RMD and RDO operations, the accuracy of observations was verified through hundreds of experiments.

4.1 Observations

Our observations indicate that the encoder does not select intra modes uniformly. Figure 3 shows the average distribution of modes, using the all intra encoding configuration [19]. According to this figure, DC and Planar modes jointly account for ~40% of all PUs in overall, which is higher than any other mode. This amount is higher for sequences with plain textures. Furthermore, it can be observed that even the distribution of angular modes is not uniform. The angular modes of 10 and 26 which are respectively horizontal and vertical modes, are considerably more frequent than the other modes. This is expected, because natural image scenes include more vertical and horizontal directions [16, 28]. Part (b) of the figure also indicates that the contribution of these modes is more significant in larger PUs, where computational complexity and consequently power consumption is considerable.

These observations are the basis of the proposed method in low power conditions: Since DC/Planar and horizontal/vertical modes are more frequent and thus more effective, the remaining angular modes can be ignored in low power conditions, especially in larger block sizes. Thus, it is logical to check the DC/Planar modes first, then check the more probable horizontal/vertical modes, and finally if the power-budget allows, try to check a few extra modes among all other angular modes.

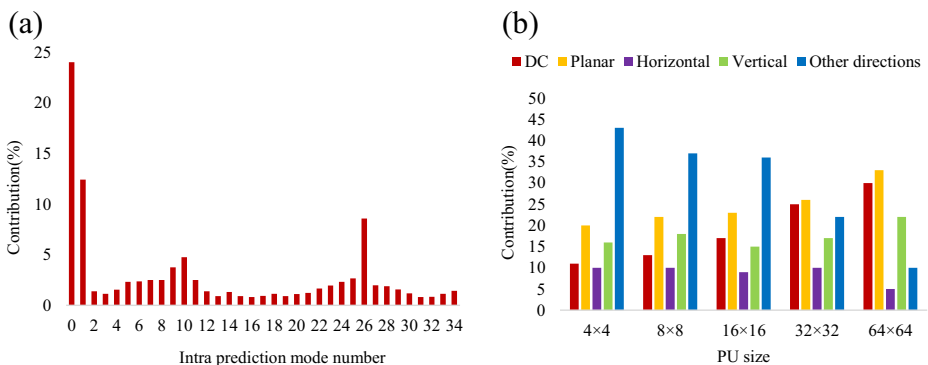


Fig. 3 a Distribution of intra prediction modes. Modes 0 and 1 are DC and Planar respectively, other modes indicate angular modes; b Distribution of intra modes in different PU sizes

Furthermore, due to their frequent occurrences, early estimation of DC/Planar modes can significantly reduce the overall coding complexity. However, the problem to be overcome is the difficulty of analyzing the DC/Planar modes. Experiments in [17] indicate that despite the common belief, these two modes are not only chosen for plain textures, but they are also used to represent regions with too much details, where all angular modes fail to find a distinct direction. Moreover, in certain sequences when the entropy of these modes is small, they might be selected despite the lower errors of other angular modes. Hence, determining these modes with texture analysis is not trivial. However, the entropy engine of HEVC predicts the required bits for each intra mode, based on recently coded blocks, which can be a good measure for estimation of these two modes. Fortunately, the experiments indicate that these two modes are the most efficient modes in modeling a random texture, and thus miss-predictions will have very limited side effect on the compression efficiency.

To analyze this side effect, ten video sequences were encoded using one of these three configurations: only DC/Planar modes, only two random angular modes, or all the 33 angular modes without DC/planar modes. The results in Table 1 show that in case of a naive prediction system where all blocks are falsely considered DC or Planar, the BD-Rate increases by 9.69%, on average. However, when all blocks are encoded by any other two random modes, BD-Rate increases by 20.41% which is twice as high as the error for DC/Planar. Even with the estimation of all the 33 angular modes, although the penalty is almost half of that of DC/Planar on average, but they are 33/2 times more complex. However, in some sequences with plain texture, such as Kimono or ParkScene, DC/Planar is even better than proper estimation of all 33 angular directions. These results indicate that if a method can estimate the selection of DC/Planar modes, it may be used without much concern, even if the estimation is sometimes not fully accurate.

4.2 Proposed intra coding process

As discussed above, the proposed fast intra prediction scheme works based on anticipating the DC and Planar modes and also detecting the texture direction. Based on the observations in Section 4.1, the overall process of the proposed method is shown in Fig. 4. In the first and second steps of the intra coding process, DC/Planar, horizontal and vertical modes are checked. Then if the power-budget can afford more angular modes, the proposed texture analysis method is used, which analyses the residual of the Planar mode to determine the dominant direction of the texture. Certain angular modes are checked according to this analysis and the available power-budget. Then the process for DC/Planar estimation might check one or two extra modes that potentially can be better than DC/Planar. If DC or Planar are proven to be

Table 1 Comparing the efficiency of DC/Planar modes with other angular modes

Class	DC and Planar modes		2 random directional modes		33 directional modes	
	BD-Rate (%)	BD-PSNR (dB)	BD-Rate (%)	BD-PSNR (dB)	BD-Rate (%)	BD-PSNR (dB)
A	15.08	-0.73	26.80	-0.71	4.52	-0.22
B	4.33	-0.16	18.43	-0.35	7.54	-0.28
C	11.95	-0.87	20.16	-1.61	3.18	-0.36
D	13.28	-0.78	16.95	-0.98	3.01	-0.18
E	6.49	-0.38	9.26	-0.54	3.37	-0.21
Avg	9.69	-0.55	20.41	-1.11	4.62	-0.25

better than those modes, the RDO process is turned off to further reduce the complexity. Otherwise, a few modes are selected as candidates for the RDO process. A flowchart of the proposed algorithm is shown in Fig. 4, and its main components are described in more detail in the following sub-sections.

4.2.1 Directional mode selection

The Intra prediction modes in HEVC, try to model the texture of a block. When the angular mode is properly selected, the residuals are smaller and the coding efficiency increases. This is achieved when the selected angular mode is oriented towards the *energy direction* of the texture. In this case, the prediction residuals will ideally carry no energy. Considering that HEVC uses DCT and zigzag scan to code the residual blocks, it can be inferred that distinct edge directions are of paramount importance in texture modeling, because they can produce sharp edges in the residuals, leading to more significant DCT coefficients and longer zigzag scans, and thus more bits. However, larger low frequency components can mask the texture through a DC shift, making it hard to discriminate the texture directions. This phenomenon can be better imagined through a simple example: assuming luminance intensities of 100, 110 and 99 in three consecutive pixels, suggests a rather flat area. But when the mean value (i.e. 103) is removed from the observations, the remainders are -3 , $+7$, and -4 , which indicate an edge. Hence, it is more efficient to first remove the plain texture and then analyze the remaining energy in horizontal and vertical directions.

While there are some signal processing tools that can provide such analysis with considerable computational and implementation overhead, in this paper and in order to minimize

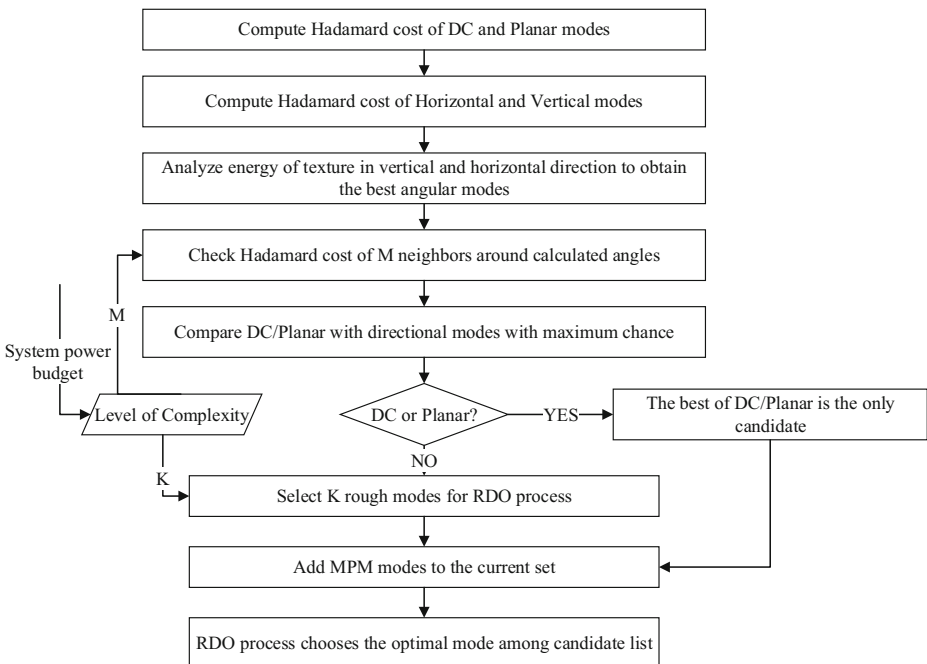


Fig. 4 The proposed algorithm for fast Intra prediction mode decision; M and K are predefined values and are set according to the level of complexity

complexity, we exploit the internal operations of the HEVC encoder itself for such analysis. The *Planar* mode, which is already an inevitable operation of the HEVC encoder, is proven to be a strong tool to model *plain* textures. Removing the predicted plain textures from the image, highlights the dominant high energy edges of the block. In other words, analyzing the residual of the Planar mode, is an efficient way of estimating the dominant edge direction. To obtain this, first the Planar filter is applied to each pixel of the block as (5) where $RefL$ and $RefT$ represent reference values from left column and top row of the block respectively and BS is the block size. Then the residual is computed as (6).

$$PlanarPred[i][j] = \frac{(BS-1-i) \times RefL(j) + (i+1) \times RefT(BS) + (BS-1-j) \times RefT(i) + (j+1) \times RefL(BS) + BS}{2} \quad (5)$$

$$PlanarRes[i][j] = OriginalBlock[i][j] - PlanarPred[i][j] \quad (6)$$

After removing the plain texture, the sharp edges will be more distinct, as the intensity of pixels changes more clearly at their vicinity. Figure 5 exemplifies this phenomenon for two random blocks from the Traffic and kimono sequences. From part (b) of the figure it can be observed that the detailed texture from the forest acts like an unpredictable source of noise which interferes with the edge analysis. The residual of the Planar mode (part (c) of the figure), includes the high energy part of the image, which contributes the most to the bitrate. The tests confirm that the selected intra mode of HEVC for these blocks (modes 8 and 21 for upper and lower images) matches the visible edge direction. After obtaining this residual, its energy in the horizontal and vertical directions is measured through (7), to estimate the texture direction.

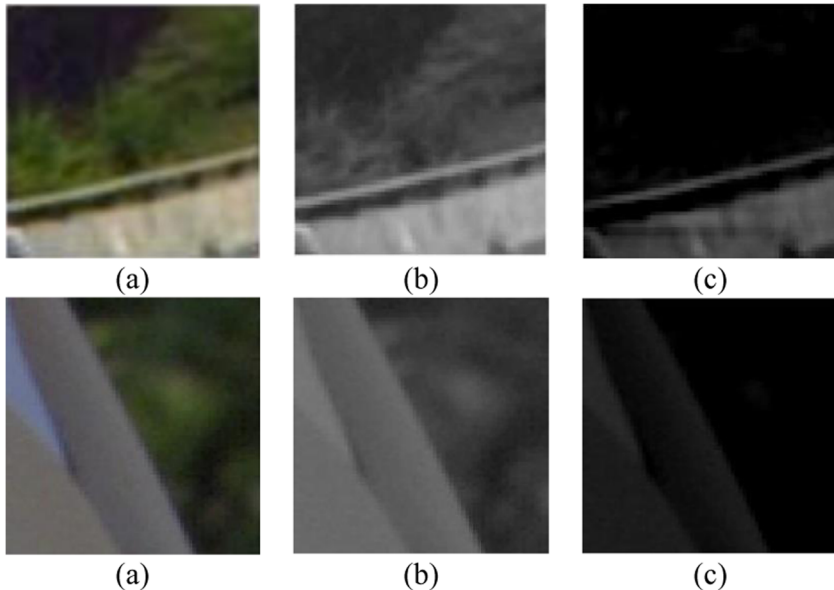


Fig. 5 Examples of edge direction analysis for two random blocks. Top row for the Traffic video and bottom for Kimono: **a** Original image, **b** Luminance of image, **c** Residual of Planar mode

$$angle = \tan^{-1} \left(\frac{\sum_{row=1}^{\#row} var_{row}}{\sum_{col=1}^{\#col} var_{col}} \right) \tag{7}$$

var_{row} and var_{col} in these equation indicate 1-D variance over row and column. Since the values of horizontal and vertical variances are positive, the *angle* is the absolute value of the texture direction. Table 2, maps each calculated direction to its corresponding intra mode. Thus, as shown in Table 2, some directions can represent two intra modes, for positive and negative directions. Based on this, the algorithm offers two estimated modes for these cases. Clearly, these modes represent only an estimate of the best direction. In order to obtain the best mode, few modes around each predicted modes should be searched, according to the available processing power. This issue is further discussed in Section 4.3.

To evaluate the applicability of this method, the test sequences were encoded using the calculated intra direction plus two extra modes around it. Also, the number of candidates for the RDO process was set to {3, 3, 2, 2, 1} to keep almost the same ratio of number of RMD to number of RDO candidates. The experimental results are tabulated in Table 3 where on average the BD-PSNR degradation and BD-Rate increase of this method compared to the baseline encoder are only 0.12 dB and 1.96%, respectively. However, the encoding time is reduced by 35.27% on average. This time reduction is achieved due to the smaller number of RMD modes and RDO candidates. It is important to note that number of RDO can be reduced, only because the proposed RMD stage is more targeted around the best direction.

4.2.2 Early DC and planar estimation

One of the built-in elements of the HEVC codec that can help in intra mode decision, is the CABAC engine, that predicts the required bits for the encoding of each intra mode. Context modeling is a key element in CABAC, which estimates the probability of each non-bypassed binary symbol based on some specific context. It selects a probability model among the 126 different tables to encode each bin value depending on the related side information (e.g. spatially left and top neighbors, depth or size of CU/PU/TU) [29]. The CABAC engine adaptively assigns bits to each intra mode, according to its probability model. According to (1) and (4), this parameter (referred to as B_{mode}) is used to measure the cost of each mode and determines the final mode. Consequently, modes with the smallest B_{mode} have the highest probability of getting selected. Since B_{mode} represents

Table 2 Angular modes and their corresponding direction

Angle	Mode(s)	Angle	Mode(s)
$\lfloor \frac{15\pi}{64}, \frac{17\pi}{64} \rfloor$	2, 18, 34	$\lfloor \frac{17\pi}{64}, \frac{19\pi}{64} \rfloor$	19, 33
$\lfloor \frac{13\pi}{64}, \frac{15\pi}{64} \rfloor$	3, 17	$\lfloor \frac{19\pi}{64}, \frac{21\pi}{64} \rfloor$	20, 32
$\lfloor \frac{11\pi}{64}, \frac{13\pi}{64} \rfloor$	4, 16	$\lfloor \frac{21\pi}{64}, \frac{23\pi}{64} \rfloor$	21, 31
$\lfloor \frac{9\pi}{64}, \frac{11\pi}{64} \rfloor$	5, 15	$\lfloor \frac{23\pi}{64}, \frac{25\pi}{64} \rfloor$	22, 30
$\lfloor \frac{7\pi}{64}, \frac{9\pi}{64} \rfloor$	6, 14	$\lfloor \frac{25\pi}{64}, \frac{27\pi}{64} \rfloor$	23, 29
$\lfloor \frac{5\pi}{64}, \frac{7\pi}{64} \rfloor$	7, 13	$\lfloor \frac{27\pi}{64}, \frac{29\pi}{64} \rfloor$	24, 28
$\lfloor \frac{3\pi}{64}, \frac{5\pi}{64} \rfloor$	8, 12	$\lfloor \frac{29\pi}{64}, \frac{31\pi}{64} \rfloor$	25, 27
$\lfloor \frac{\pi}{64}, \frac{3\pi}{64} \rfloor$	9, 11	$\lfloor \frac{31\pi}{64}, \frac{\pi}{2} \rfloor$	26
$\lfloor 0, \frac{\pi}{64} \rfloor$	10	–	–

Table 3 Encoding efficiency and complexity reduction of the proposed algorithm in mode decision compared to HM

Class	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)
A	2.11	-0.11	35.62
B	0.76	-0.02	35.90
C	2.41	-0.19	34.19
D	3.61	-0.21	34.53
E	1.51	-0.08	35.76
avg	1.96	-0.12	35.27

a rough probability of each intra mode, this paper uses this parameter to reduce the RDO processing time whenever DC/Planar are detected.

Our investigations indicate that when the Hadamard cost of DC or Planar modes is less than the modes with the smallest coding bits, most probably the RDO process chooses the DC or Planar modes as the final winner. Thus, when such condition is met, the proposed method sets the best of DC/Planar as the final intra mode and skips the RDO process. Table 4 presents the accuracy of this method. In this experiment, ten video sequences were encoded with four different QPs, and the modes with the smallest B_{mode} were used to determine the DC/Planar modes and the results were compared with the baseline encoding process. The experimental results indicate that in ~90% of cases when the baseline encoder encodes a PU with DC/Planar modes, the proposed method can detect DC/Planar modes correctly.

False positives in this method occur when there are sudden and unexpected changes in the texture. In this case, the CABAC engine which assigns bits based on the previously coded textures, still uses a large B_{mode} for specific directions in the new region. This leads to ignoring those directions. However, it should be noted that in this situation DC/Planar still have a smaller cost than the modes with the smallest B_{mode} , and thus the newly introduced directions are usually marginally better. Moreover, after such an incident, CABAC will update its texture models and thus the B_{mode} of next blocks will be accurate again.

Extensive experiments confirm the high accuracy of this method as depicted in Table 5. To measure the pure effect of this algorithm, in this experiment all angular intra modes were checked in the RMD step. Then, when the cost of the best of DC/Planar is better than those

Table 4 Accuracy (%) that the final optimal mode by RDO is planar or DC mode when Hadamard cost of DC or Planar modes is less than that of modes with smallest coding bits

Class	Sequence	QP			
		22	27	32	37
A	PeopleOnStreet	88.46	87.64	84.85	88.66
	Traffic	88.16	90.84	87.79	85.71
B	Kimono	92.64	91.58	92.20	91.27
	ParkScene	84.10	87.23	85.86	86.05
C	PartyScene	93.19	93.22	92.56	90.28
	RaceHorses480	90.82	91.38	88.20	87.67
D	BlowingBubble	86.27	92.00	94.33	91.67
	RaceHorses	92.85	92.98	88.70	92.98
E	City	88.75	88.77	88.00	88.05
	MobileCalendar	89.75	95.57	98.28	99.11
avg		89.50	91.12	90.08	90.14

Table 5 Encoding efficiency and complexity reduction of the proposed algorithm in early DC and Planar estimation compared to HM

Class	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)
A	0.89	-0.04	21.31
B	0.49	-0.01	25.12
C	0.96	-0.07	22.78
D	1.09	-0.06	22.26
E	0.61	-0.03	22.51
avg	0.78	-0.04	23.01

modes which their B_{mode} are smallest among angular modes, the RDO process was skipped. As the results indicate, the proposed method can reduce the total encoding time by 23%, while causing only 0.78% increase in bitrate.

Combining the DC/Planar estimation and the directional mode selection schemes in Section 4.2.1 and 4.2.2, the overall performance of the proposed fast mode decision scheme is provided in Table 6. As the results indicate, the proposed method achieves 40.9% saving in total encoding time with only 0.11 dB BD-PSNR degradation and 1.92% BD-Rate increase with respect to HM. The time reduction of the combined method is more significant than both the methods separately. But since there is a subset that is common in both methods, their combination does not lead to summation of their time reductions. This is because the RDO process is already limited in the directional mode selection scheme, thus some parts of the time reduction from the two methods overlap.

4.3 Proposed computationally scalable intra coding scheme

Exploiting the statistical information provided in Section 4.1, a computationally scalable intra coding scheme can be proposed, such that the different all intra configurations of this scheme can be used for different levels of processing power.

The proposed scheme consists of seven predefined levels of complexity, with nearly equal distances in term of complexity. The complexity difference between two neighboring configurations in the proposed method is designed to be ~5% to uniformly cover the real case scenarios in medium grain complexity control. Furthermore, to have the best possible compression efficiency within each complexity level, the number of RMD and RDO processes in each level were obtained through extensive experiments on real case video sequences. Table 7 lists these complexity levels:

Table 6 Overall encoding efficiency and complexity reduction of the proposed fast mode decision algorithm compared to HM

Class	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)
A	2.13	-0.11	39.97
B	0.87	-0.03	43.12
C	2.36	-0.18	39.72
D	3.31	-0.19	39.63
E	1.47	-0.08	40.94
avg	1.92	-0.11	40.90

Table 7 Parameter setting for different complexity levels of the proposed method

Level of complexity	Modes in RMD process	RDO process for $\{4 \times 4, 8 \times 8, 16 \times 16, 32 \times 32, 64 \times 64\}$
Level 1	2 modes (DC + Planar)	{1, 1, 1, 1}
Level 2	4 modes (DC + Planar+H + V)	{2, 2, 1, 1}
Level 3	10 modes (DC + Planar+H + V + 6 directional mode) + DC & Planar estimation	{3, 3, 2, 1}
Level 4	14 modes (DC + Planar+H + V + 10 directional mode) + DC & Planar estimation	{4, 4, 2, 1}
Level 5	18 modes (DC + Planar+H + V + 14 directional mode) + DC & Planar estimation	{6, 6, 3, 3, 1}
Level 6	26 modes (DC + Planar+H + V + 22 directional mode) + DC & Planar estimation	{7, 7, 3, 3, 1}
Level 7	35 modes + DC & Planar estimation	{8, 8, 3, 3, 3}

Level 1 is dedicated to the lowest computational complexity and consequently results in the lowest encoding quality. It is suitable when the device is extremely low power and can afford a minimum level of processing. Based on observations in 4.1, in Level 1, only the DC and Planar modes are considered. Also, since only two modes are being processed, there is no need for RDO to decide the best mode, hence it is turned off. Level 1 is the minimum acceptable power. A device with a power less than this amount is considered incapable of HEVC encoding

In Level 2, when the device is capable of providing more processing power, the horizontal and vertical modes will also be checked with this extra power. Also, RDO is turned on for small block sizes to decide among these four modes. However, since there are only four modes, it is experimentally observed that only two candidates for the RDO process are enough. It is important to note that for large block sizes, RDO has a very limited impact on the coding efficiency. In the extreme case of 64×64 pixels blocks, the RDO process has the slightest impact on the final decision and experiments confirm that the best mode in RMD, is almost always the best mode after RDO. The importance of RDO increases for smaller block sizes.

In Level 3, according to the proposed method in 4.2.1, one or two angular modes (corresponding to positive and negative directions) will be estimated. Then, these modes, plus one mode around them (meaning angle1 -1 to angle1 + 1, and angle2 -1 to angle2 + 1) are checked. Also, the number of RDO candidates is increased compared to the previous level. This number is derived based on experiments, according to how many modes are checked in the RMD process.

In Levels 4 to 6, more directional modes are tested around each estimated direction in RMD. Number of RDO candidates also increases to improve the compression quality. In Level 4, angle -2 and angle +2 are added to the Level 3 mode set; resulting in the calculated directions and four neighbors around them. In Level5 and Level 6, respectively six (from angle -3 to angle +3) and ten (from angle -5 to angle +5) neighbors are checked in the RMD process. Levels 4 to 6, can be considered as high quality and low power fast intra coding schemes for conventional devices.

Finally, level 7 is considered for devices with high end processing power, where the designer is willing to spend more computations for better coding efficiency. In this configuration, all 33 intra modes are checked during the RMD process and only the

DC/Planar decision in 4.2.2 is used to reduce the number of RDO candidates when DC/Planar are detected. This Level provides excellent coding efficiency while still saving considerable computational complexity and thus is considered as a semi-low power high quality encoder. For devices with higher processing power, where the coding efficiency is extremely important, the baseline encoder is suggested. Although it should be noted that such decision is extremely rare, firstly because of the high processing power, and secondly because the coding efficiency of Level 7 is already excellent and does not have a noticeable difference with the baseline encoder.

5 Experimental results

To evaluate the performance of the proposed method, it has been implemented on top of the HEVC reference software HM16.14 [19]. The tests have been performed on a @2.80 GHz Intel Core i7 CPU with 8-GBs of memory. All experiments have been repeated for the proposed method and the baseline HM encoder, using the same all-intra configuration and common test conditions [2]. In the experiments, the performance is evaluated for 11 video sequences of five classes with different resolutions (all required resolutions by the common test conditions) in the YUV 4:2:0 format. Table 8 lists the 11 video sequences used in this paper with more detailed information. All sequences were encoded using each of the seven computational complexity levels given in Table 7, using QPs 22, 27, 32, and 37, and BD-Rate and BD-PSNR were used to compare the results with corresponding results from HM encoder.

Table 9 tabulates the performance of the proposed method in each complexity level. As it can be verified from Table 9, the seven predefined levels of complexity provide intra coding with 23 to 52% total time reductions compared to the baseline HEVC encoder, with respectively 0.78 to 9.69% increase in BD-Rate. The complexity of each level increases by about 5% to uniformly cover the range from 23 to 52%, which is acceptable for a time control of medium granularity [6].

Based on the analysis in Section 2, the intra prediction alongside the RDO process accounts for nearly 55% of the total encoding time. Thus the 52% time saving is nearly the maximum possible amount at Level 1. Although BD-Rate of this level is noticeable (9.6%), use of DC

Table 8 Detailed information of the test sequences

Class	Sequence	Resolution	Frame Rate
A	PeopleOnStreet	2560 × 1600	30
	Traffic		30
B	Kimono	1920 × 1080	24
	ParkScene		24
	SunFlower		25
C	PartyScene	832 × 480	50
	RaceHorses480		30
D	BlowingBubble	416 × 240	50
	RaceHorses		30
E	City	1280 × 720	60
	MobileCalendar		50

Table 9 Encoding efficiency and complexity reduction of the proposed algorithm compared to HM

Class	LEVEL1			LEVEL2			LEVEL3			LEVEL4		
	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)
A	15.08	-0.73	51.85	10.59	-0.52	46.40	2.13	-0.10	39.97	2.07	-0.10	35.61
B	4.33	-0.17	54.05	3.21	-0.12	48.83	0.86	-0.03	43.12	0.79	-0.03	39.70
C	11.95	-0.87	50.45	8.07	-0.61	45.36	2.35	-0.18	39.72	1.96	-0.15	35.21
D	13.28	-0.78	50.55	10.27	-0.61	45.30	3.31	-0.20	39.63	2.51	-0.15	34.51
E	6.49	-0.38	51.51	4.129	-0.24	46.32	1.48	-0.09	40.94	1.23	-0.07	36.27
avg	9.69	-0.55	51.90	6.89	-0.39	46.66	1.92	-0.11	40.90	1.63	-0.09	36.579
(a)												
Class	LEVEL5			LEVEL6			LEVEL7					
	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)			
A	1.48	-0.07	30.00	1.18	-0.06	25.63	0.89	-0.04	21.31			
B	0.63	-0.02	35.30	0.54	-0.02	31.14	0.49	-0.02	25.12			
C	1.66	-0.13	30.62	1.27	-0.10	26.44	0.96	-0.07	22.78			
D	2.07	-0.12	29.95	1.57	-0.09	25.48	1.09	-0.06	22.26			
E	0.99	-0.06	30.73	0.71	-0.04	26.04	0.60	-0.03	22.50			
avg	1.30	-0.07	31.68	1.01	-0.06	27.33	0.78	-0.04	23.01			
(b)												

and Planar guarantees the best possible coding efficiency within this level of power budget. Also it should be noted that in extremely low levels of processing power, it might be impossible for a device to perform real time encoding even with a low quality. Therefore, the main goal of this method is to enable HEVC encoding in such situation, with acceptable coding efficiency.

With ~5% more power, in level 2, adding horizontal and vertical directions improves BD-Rate by 2.8%. With another ~5% more power, the method adds the calculated direction to the RMD process which improves BD-Rate by 5%. Adding extra angular modes for further refinement, BD-Rate improves in levels 4 to 7 around 0.3% in each level compared with its previous level, while the encoding time continues to increase with a nearly steady pace.

In Table 10, the proposed method is compared with the fast intra coding technique of [34] which is one of the state-of-the-art techniques. Since Level 3 of the proposed method has a comparable coding efficiency with this method, it is selected for comparison. It can be observed that the proposed method is 6.44% faster with almost the same coding efficiency. Given that both methods use texture analysis techniques, the superiority of the proposed method is due to the filtering with near zero overhead, and the early termination of RDO for DC/Planar modes.

To have a better perspective of the scalability of the proposed method, a more abstract comparison in terms of complexity reduction and compression efficiency with several fast intra coding schemes is provided in Fig. 6. The red line shows the proposed scalable fast intra coding with seven levels of complexity that can provide 20%-50% improvement in computational complexity. In contrast, each of the other fast methods (shown with black circles) can provide a single solution, suitable for a narrow range power budgets. Although some methods in Fig. 6 may have a slightly better performance than the proposed method, it should be considered that they cannot be expanded and cover the complexity domain. The figure indicates that different complexity levels of this method are either better or at least comparable with the fast methods at similar encoding times. As shown in this figure, in the low power situations (right side of the figure) when it needs more than 40% complexity reduction, the proposed method makes it possible to encode with the minimum processing power, while no other method can provide such opportunity through intra mode decision. However, it should be noticed that this complexity reduction can be achieved through early CU size decision which is out of the scope of this paper, and also can be used alongside the proposed method to improve the time saving. At the high power end of the figure (left side), the proposed method provides excellent coding efficiency with better encoding time compared to similar methods. This valuable feature makes this method suitable for almost all kind of implementations, from high end powerful encoder boards, to extremely low power surveillance cameras.

Table 10 Performance of Level 3 compared against the technique in [34]

Class	BD-Rate (%)	BD-PSNR (dB)	Enc. Time Reduction (%)
A	0.21	0.03	-2.32
B	-0.11	0.01	-2.70
C	0.29	-0.03	-12.9
D	0.55	-0.02	-7.74
E	-	-	-
avg	0.23	0.00	-6.44

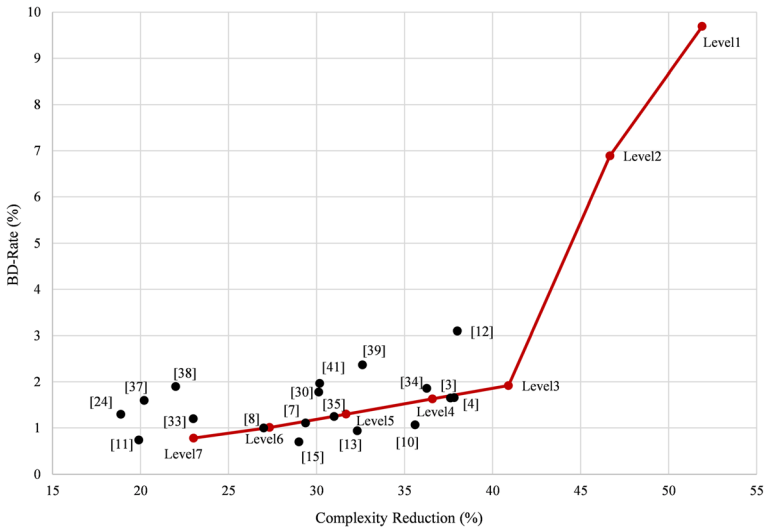


Fig. 6 Comparing related works with the proposed computationally scalable method

6 Conclusion

This paper presented a fast intra coding scheme which is computationally scalable. The paper reveals that Planar, DC, vertical and horizontal intra modes are respectively the most frequently used modes and constitute the majority of intra coded blocks. Thus, extremely low power devices can perform encoding by only checking these modes. Also exploiting the efficiency of the Planar mode in modeling the plain texture, the paper shows that the direction of the texture can be detected early-on by analyzing the residual of this mode. Moreover, the proposed method can detect the DC/Planar modes with the help of CABAC entropy encoder at almost no extra complexity, which leads to further reduction of the RDO time. The flexibility of the proposed method in terms of number of tested modes, provides the opportunity of introducing seven complexity levels for various electronic devices. Experiments confirm that the proposed method is comparable to the state-of-the-art techniques at different levels of complexity, with total encoding time reductions from ~23% up to ~52%, and thus can be employed by many different implementations.

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

References

1. Bjontegaard G (2001) Calculation of average PSNR differences between RD-curves. ITU-T Q6/SG16, Doc VCEG-M33, 15th Meet Austin, Texas
2. Bossen F (2011) Common test conditions and software reference configurations. Jt Collab Team Video Coding ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, 5th Meet
3. Chen G, Liu Z, Ikenaga T, Wang D (2013) Fast HEVC intra mode decision using matching edge detector and kernel density estimation alike histogram generation. Proc IEEE Int Symp Circuits Syst 53–56. <https://doi.org/10.1109/ISCAS.2013.6571780>

4. Chung B, Kim J, Yim C (2014) Fast rough mode decision method based on edge detection for intra coding in HEVC. In: The 18th IEEE International Symposium on Consumer Electronics (ISCE 2014). IEEE, pp 1–2
5. Correa G, Assuncao P, Agostini L, da Silva Cruz LA (2012) Performance and computational complexity assessment of high-efficiency video encoders. *IEEE Trans Circuits Syst Video Technol* 22:1899–1909. <https://doi.org/10.1109/TCSVT.2012.2223411>
6. Correa G, Assuncao PA, Agostini LV, Da Silva Cruz LA (2016) Pareto-based method for high efficiency video coding with limited encoding time. *IEEE Trans Circuits Syst Video Technol* 26:1734–1745. <https://doi.org/10.1109/TCSVT.2015.2469533>
7. Ding W, Shen W, Shi Y, Yin B (2014) A fast intra-mode decision scheme for HEVC. In: 2014 5th International Conference on Digital Home. IEEE, pp 70–73
8. Gao L, Dong S, Wang W, et al (2015) Fast intra mode decision algorithm based on refinement in HEVC. In: 2015 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, pp 517–520
9. Ha JM, Bae JH, Sunwoo MH (2016) Texture-based fast CU size decision algorithm for HEVC intra coding. In: 2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS). IEEE, pp 702–705
10. Jamali M, Coulombe S, Caron F (2015) Fast HEVC intra mode decision based on edge detection and SATD costs classification. In: 2015 Data Compression Conference. IEEE, pp 43–52
11. Jiang W, Ma H, Chen Y (2012) Gradient based fast mode decision algorithm for intra prediction in HEVC. In: 2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet). IEEE, pp 1836–1840
12. Kim Y, Jun D, Jung SH et al (2013) A fast intra-prediction method in HEVC using rate-distortion estimation based on hadamard transform. *ETRI J* 35:270–280. <https://doi.org/10.4218/etrij.13.0112.0223>
13. Li F, Jiao D, Shi G et al (2017) An AR based fast mode decision for H.265/HEVC intra coding. *Multimed Tools Appl* 76:13107–13125. <https://doi.org/10.1007/s11042-016-3737-z>
14. Lin Y-C, Lai J-C, Cheng H-C (2016) Coding unit partition prediction technique for fast video encoding in HEVC. *Multimed Tools Appl* 75:9861–9884. <https://doi.org/10.1007/s11042-015-2778-z>
15. Liu X, Liu Y, Wang P et al (2017) An adaptive mode decision algorithm based on video texture characteristics for HEVC intra prediction. *IEEE Trans Circuits Syst Video Technol* 27:1737–1748. <https://doi.org/10.1109/TCSVT.2016.2556278>
16. Min J-H, Lee S, Kim I-K et al (2010) Unification of the directional intra prediction methods in TMuC. Document: JCTVC-B100, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Geneva
17. Pakdaman F, Hashemi M-R, Ghanbari M (2017) Fast and efficient intra mode decision for HEVC, based on dual-tree complex wavelet. *Multimed Tools Appl* 76:9891–9906. <https://doi.org/10.1007/s11042-016-3584-y>
18. Ramezanzpour Fini M, Zargari F (2016) Two stage fast mode decision algorithm for intra prediction in HEVC. *Multimed Tools Appl* 75:7541–7558. <https://doi.org/10.1007/s11042-015-2675-5>
19. Rosewarne C, Bross B, Naccari M, Sharman K (2015) High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Encoder Description. Document: JCTVC-V1002, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC), Geneva
20. Ruiz D, Fernández-Escribano G, Adzic V et al (2017) Fast CU partitioning algorithm for HEVC intra coding using data mining. *Multimed Tools Appl* 76:861–894. <https://doi.org/10.1007/s11042-015-3014-6>
21. Semsarzadeh M, Lotfi A, Hashemi MR, Shirmohammadi S (2013) A fine-grain distortion and complexity aware parameter tuning model for the H.264/AVC encoder. *Signal Process Image Commun* 28:441–457. <https://doi.org/10.1016/j.image.2013.02.003>
22. Sharabayko MP, Markov NG (2014) Entropy-based intra-coding RDO estimation for HEVC. 2014 9th Int Forum Strateg Technol IFOST 2014 56–59. <https://doi.org/10.1109/IFOST.2014.6991071>
23. Shen L, Zhang Z, An P (2013) Fast CU size decision and mode decision algorithm for HEVC intra coding. *IEEE Trans Consum Electron* 59:207–213. <https://doi.org/10.1109/TCE.2013.6490261>
24. Silva TL, Agostini LV, Luis A, Cruz S (2012) Fast HEVC intra prediction mode decision based on EDGE direction information. In: Proc. 20th Eur. Signal Process. Conf. (EUSIPCO). pp 1214–1218
25. da Silva TL, Agostini LV, da Silva Cruz LA (2015) Fast mode selection algorithm based on texture analysis for 3D-HEVC intra prediction. In: 2015 IEEE International Conference on Multimedia and Expo (ICME). IEEE, pp 1–6
26. Song Y, Zeng Y, Li X et al (2017) Fast CU size decision and mode decision algorithm for intra prediction in HEVC. *Multimed Tools Appl* 76:2001–2017. <https://doi.org/10.1007/s11042-015-3155-7>
27. Sullivan GJ, Ohm J-R, Han W-J, Wiegand T (2012) Overview of the High Efficiency Video Coding (HEVC) standard. *IEEE Trans Circuits Syst Video Technol* 22:1649–1668. <https://doi.org/10.1109/TCSVT.2012.2221191>
28. Sze V, Madhukar Budagavi GJS (2014) High Efficiency Video Coding (HEVC). Springer
29. Sze V, Marpe D (2014) Entropy coding in HEVC. Springer, Cham, pp 209–274

30. Tian R, Zhang Y, Fan R, Wang G (2016) Adaptive fast mode decision for HEVC intra coding. In: 2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA). IEEE, pp 1–6
31. Vanne J, Viitanen M, Hamalainen TD, Hallapuro A (2012) Comparative rate-distortion-complexity analysis of HEVC and AVC video codecs. *IEEE Trans Circuits Syst Video Technol* 22:1885–1898. <https://doi.org/10.1109/TCSVT.2012.2223013>
32. Wiegand T, Sullivan G (2003) Joint Video Team (JVT) of ISO / IEC MPEG & ITU-T VCEG Document: JVT-G050 Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec . H . 264 | ISO / IEC 14496-10 AVC) Approved Output Document of JVT Tex. 2002:7–14
33. Yan S, Hong L, He W, Wang Q (2012) Group-based fast mode decision algorithm for intra prediction in HEVC. 8th Int Conf Signal Image Technol Internet Based Syst SITIS 2012r 225–229. <https://doi.org/10.1109/SITIS.2012.41>
34. Yao Y, Li X, Lu Y (2016) Fast intra mode decision algorithm for HEVC based on dominant edge assent distribution. *Multimed Tools Appl* 75:1963–1981. <https://doi.org/10.1007/s11042-014-2382-7>
35. Yao F, Zhang X, Gao Z, Yang B (2016) Fast mode and depth decision algorithm for HEVC intra coding based on characteristics of coding bits. In: 2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB). IEEE, pp 1–4
36. Zhang H, Ma Z (2012) Fast intra prediction for high efficiency video coding. Springer, Berlin, pp 568–577
37. Zhang M, Zhao C, Xu J (2012) An adaptive fast intra mode decision in HEVC. *Proc Int Conf Image Process ICIP* 221–224. <https://doi.org/10.1109/ICIP.2012.6466835>
38. Zhang H, Liu Q, Ma Z (2014) Fast intra mode decision for high efficiency video coding. *IEEE Trans Circuits Syst Video Technol* 24:660–668. <https://doi.org/10.1109/PCS.2013.6737739>
39. Zhao W, Onoye T, Song T (2015) Hierarchical structure-based fast mode decision for H.265/HEVC. *IEEE Trans Circuits Syst Video Technol* 25:1651–1664. <https://doi.org/10.1109/TCSVT.2015.2395751>
40. Zhu S, Zhang C (2017) A fast algorithm of intra prediction modes pruning for HEVC based on decision trees and a new three-step search. *Multimed Tools Appl* 76:21707–21728. <https://doi.org/10.1007/s11042-016-4056-0>
41. Zhu S, Zhao D, Li L (2015) Adaptive fast intra prediction for high efficiency video coding. *Multimed Tools Appl* 75:759–7573. <https://doi.org/10.1007/s11042-015-2677-3>



Elahe Hosseini received her B.Sc. degree in Computer Engineering from Amirkabir University of Technology. She is currently pursuing her M.Sc. degree in computer engineering at the University of Tehran. Her research at the Multimedia Processing Lab (MPL) includes power-aware and power-efficient video encoding systems.



Farhad Pakdaman received his B.Sc. and M.Sc. degrees both in computer engineering from University of Mazandaran and the University of Tehran in 2011 and 2013, respectively. He is currently pursuing his Ph.D. in Computer Engineering at the University of Tehran with focus on low complexity and power aware algorithms and architectures for HEVC video encoding. His research interests include multimedia systems and video coding, parallel processing, computer architecture, high performance computing, and application specific architectures. From 2014 to 2017 he has served as a lecturer at the University of Mazandaran.



Mahmoud Reza Hashemi received his B.Sc., and M.Sc. in Electrical Engineering from the University of Tehran. He pursued his Ph.D. at the University of Ottawa, Canada. He is currently an Associate Professor at the School of Electrical and Computer Engineering of the University of Tehran. Dr. Hashemi is the co-founder and Director of the Multimedia Processing Laboratory (MPL). His research interests include reconfigurable hardware architectures for multimedia processing, computer architecture, multimedia systems and networking, and recently cloud gaming. Dr. Hashemi is a member of the editorial board of several journals, and has chaired or served on the organizing and program committees of numerous international conferences in multimedia, and information technology.



Mohammad Ghanbari (M'78, SM'97, F'01, LF14) is a professor of video networking at the University of Tehran, Iran and Emeritus Professor at the university of Essex, United Kingdom. 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 thirteen 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 to Life Fellow of IEEE.