Fast Intra Mode Decision for HEVC

Kanayah Saurty, Pierre C. Catherine and Krishnaraj M. S. Soyjaudah

Abstract High Efficiency Video Coding (HEVC/H.265), the latest standard in video compression, aims to halve the bitrate while maintaining the same quality or to achieve the same bitrate with an improved quality compared to its predecessor, AVC/H.264. However, the increase in prediction modes in HEVC significantly impacts on the encoder complexity. Intra prediction methods indeed iterate among 35 modes for each Prediction Unit (PU) to select the most optimal one. This mode decision procedure which consumes around 78 % of the time spent in intra prediction consists of the Rough Mode Decision (RMD), the simplified Rate Distortion Optimisation (RDO) and the full RDO processes. In this chapter considerable time reduction is achieved by using techniques that use fewer modes in both the RMD and the simplified RDO processes. Experimental results show that the average time savings of the proposed method indeed yields a 42.1 % time savings on average with an acceptable drop of 0.075 dB in PSNR and a negligible increase of 0.27 % in bitrate.

1 Introduction

High Efficiency Video Coding (HEVC) or H.265, developed by ISO-IEC/MPEG and ITU-T/VCEG which formed the Joint Collaborative Team on Video Coding (JCT-VC) [11] became the new standard in video compression technologies in 2013. It is an evolution of the block-based hybrid motion-compensation and transform coding inherent in the previous standard, AVC/H.264. Coding performance by almost twofold is achieved by the new standard through the implementation of additional

K. Saurty
Université des Mascareignes, Pamplemousses, Mauritius
e-mail: ksaurty@udm.ac.mu

P.C. Catherine (✉)
University of Technology, La Tour Koenig, Mauritius
e-mail: ccatherine@umail.utm.ac.mu

K.M.S. Soyjaudah (✉)
University of Mauritius, Réduit, Mauritius
e-mail: ssoyjaudah@uom.ac.mu

Y. Bi et al. (eds.), *Intelligent Systems and Applications*,
Studies in Computational Intelligence 650, DOI 10.1007/978-3-319-33386-1_18

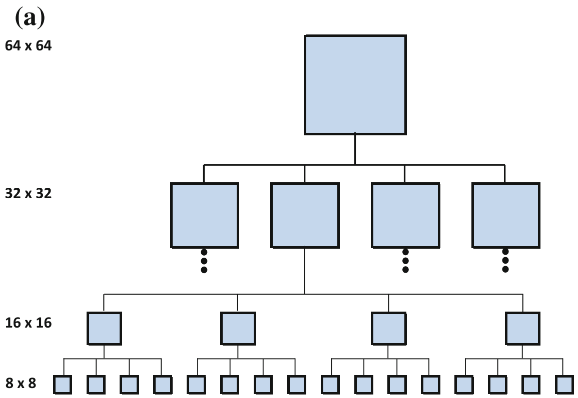

Fig. 1 **a** HEVC quad-tree structure. **b** Example of a final LCU structure

coding tools such as the newly introduced recursive tree structure, larger block transforms, adaptable loop filtering operations and advanced motion prediction [21]. In fact, video compression under HEVC aims at only half of bitrate required by the previous standard for the same quality of pictures.

The support of Coding Tree Block (CTB) of larger size compared to the conventional 16 × 16 macroblocks in H.264 is one of the major improvements in HEVC. The Largest Coding Unit (LCU), into which a frame is partitioned in HEVC, can adopt a size of either 16 × 16, 32 × 32 or 64 × 64 depending on the encoder configuration. This large sized LCU can be recursively split forming a quad-tree structure allowing sub-CUs with size of 64 × 64, 32 × 32, 16 × 16 and 8 × 8 as shown in Fig. 1a. In addition, CUs are further split into Prediction Units (PUs). In intra prediction, each CU can have a single $2N \times 2N$ PU or it can be partitioned into four $N \times N$ PUs, where N is half of the CU size. The leaf CUs in Fig. 1a may therefore be again split into 4 × 4 PUs. Depending on the texture of the picture, these LCUs will have many possible configurations. A large homogeneous region may be captured by a single CU of a large size which can be encoded using a smaller number of symbols compared to several smaller blocks. An example of a final LCU structure is shown in Fig. 1b illustrating different CU size combinations within a single LCU.

Support for the arbitrary Largest Coding Tree Block (LCTB) of different sizes enables the codec to be readily optimized for various contents, applications and devices. By choosing the appropriate LCTB size and the maximum hierarchical depth, the block structure can be optimized in a better way for the targeted application [17]. Another improvement in the new standard is the large number of intra prediction modes at the level of PUs. Compared to the 9 modes in AVC [23], HEVC makes use of 35 modes consisting of 33 directional modes along with the Intra Planar and the DC modes. This is illustrated in Fig. 2. The large number of prediction modes results in a more accurate intra prediction which consequently leads to smaller residuals that need to be processed. The increase in prediction modes therefore contributes to the high quality of the compression along with a lower bitrate.

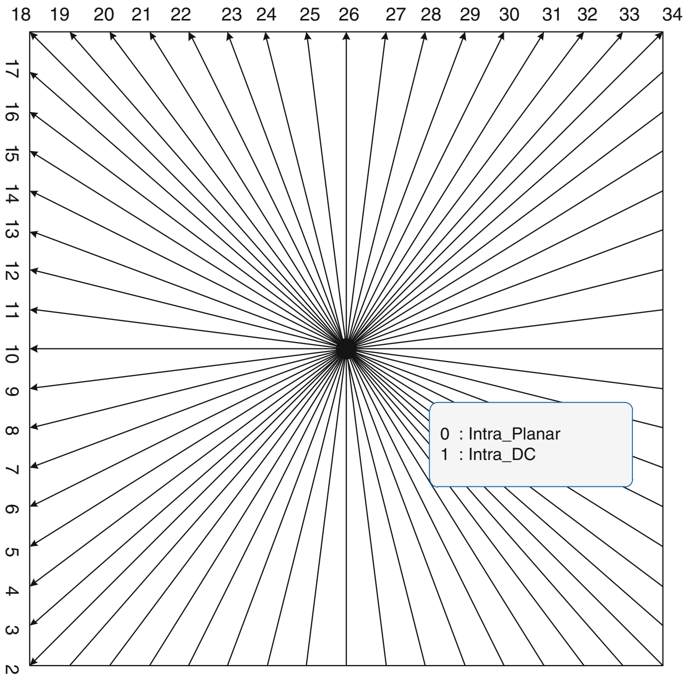

Fig. 2 HEVC intra prediction modes [10]

However, an increase in prediction modes also translates to more comparisons to be performed in order to come up with the best mode. Therefore, the benefit of higher quality derived from the increased in number of modes comes along with an increase in the encoder complexity. It therefore requires an appropriate mode selection heuristics in order to maintain a reasonable search complexity [2]. Consequently, for practical applications such as high-resolution video services and real-time processing, HEVC still requires an acceptable reduction in complexity while maintaining the high coding performance [22].

To reduce the intra coding complexity, many techniques have been proposed in other works. For example, [6] proposes an intra mode decision based on dominant edge evaluation and tree structure dependencies which yields a 20 % time savings along with an increase in bitrate of 0.9 % and a decrease of 0.02 dB in PSNR. A gradient based fast mode decision algorithm proposed in [12] results in a 20 % decrease in encoding time and a drop of 0.04 dB BD-PSNR together with a BD-Rate of 0.74 %. A decrease of 19 % in complexity is proposed by [20] at the expense of a 0.4 % increase in BD-Rate. In [14], the number of candidates for the given block are reduced by checking the boundary pixels resulting in a 13 % time savings accompanied by a difference of −0.05 dB in terms of PSNR and an increase of 1.78 % for the bitrate. A 28 % reduction is reached according to [4] together with a BD-Rate of 0.53 % and a BD-PSNR of −0.038 dB.

Table 1 Performances of other works using reduced mode

No	Algorithm	Time (%)	Δ PSNR (dB)	Δ Bit Rate (%)	Δ BD-PSNR (dB)	Δ BD- Rate (%)
1	Kumar V. [14]	−13	−0.05	1.78	-	-
2	Yongfang Shi [20]	−19	-	-	-	0.40
3	Wei Jiang [12]	−20	-	-	−0.04	0.74
4	da Silva [6]	−20	−0.02	0.9	-	1.3
5	Gaoxing Chen [4]	−28	-	-	−0.038	0.53
6	da Silva [5]	−37	−0.07	-	-	1.2
7	Guang Chen [3]	−37.61	-	-	−0.078	1.65
8	Saurty K. [19]	−38.0	−0.059	0.30	−0.063	1.15

A complexity reduction of 37.61 % is achieved in [3] with a BD-PSNR of −0.0781 dB and a BD-Rate of 1.65 %. The 37 % time savings proposed by [5] is accompanied by a −0.07 dB drop in PSNR and a BD-Rate of 1.2 %. In [19], a reduced mode approach based on the analysis of the graph obtained by plotting the HAD cost of the modes against their mode number results in 38.0 % time savings on average along with a decrease in PSNR of 0.059 dB and a 0.30 % increase in bitrate. A summary of the performances these works is also provided in Table 1. A hyphen (-) denotes data not available from the source.

In order to improve further the complexity reduction in HEVC intra prediction, a number of approaches have been proposed. They focus on the reduction in modes processed by the simplified RDO in addition to using fewer modes in the RMD process. The rest of this chapter is thus organized as follows: Sect. 2 provides an overview of intra mode decision in HEVC. Section 3 illustrates the different time-consuming components of HEVC intra prediction along with the mode selection frequency. Section 4 looks into the methods to reduce the modes taking part in the RMD process. The different approaches to reduce the number of modes entering the simplified RDO process are explained in Sect. 5. Finally, Sect. 6 concludes this chapter.

2 Intra Mode Decision in HEVC

Intra prediction in HEVC makes use of 33 angular modes along with the DC and the Planar modes for each PU [10] as shown in Fig. 2. For each PU, intra prediction is performed by extrapolating the reference pixels, obtained from the already decoded neighboring PUs, in the prediction directions for the angular modes. An example of the extrapolation for the directional mode 29 [21] is shown in Fig. 3.

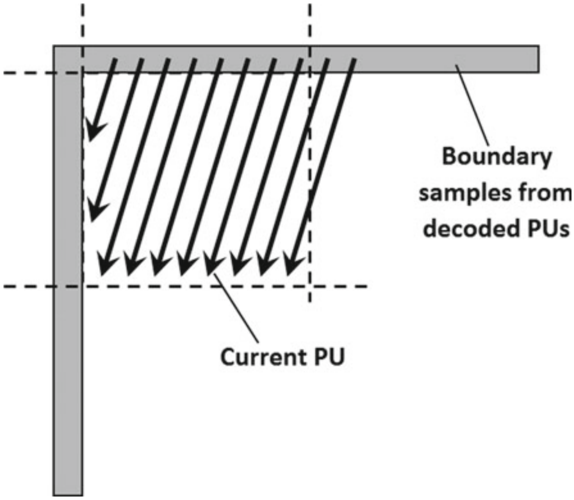

Fig. 3 Directional mode 29 for intra prediction

Table 2 Intra prediction modes combinations for a 64 × 64 CU

Size	Number of PUs	Number of modes for each PU	Total number of modes for each level
64 × 64	1	35	35
32 × 32	4	35	140
16 × 16	16	35	560
8 × 8	64	35	2,240
4 × 4	256	35	8,960
Total			11,935

The number of arithmetic operations to calculate the prediction for even one mode is quite complex and time consuming. Computing the full RDO for all of these modes in order to choose the best one would therefore require extensive and unacceptable computing resources. This is illustrated in Table 2 where the total number of possible mode combinations for a 64 × 64 CU is 11,935.

Since this process is very time consuming, HEVC already adopts a simplified three steps approach [18] which is illustrated in Fig. 4.

1. A reduced set of N candidate modes is first identified through a RMD process based on the Hadamard transform Absolute Difference (HAD) and the estimated bitrate for the various modes (35 in all). The Most Probable Modes (MPMs) obtained from the decoded neighboring PUs are added to the N candidates if they are not already included in the list.
2. A simplified Rate Distortion (RD) cost is then computed for each of the selected candidates in the previous step and the best mode identified.
3. Finally the full RD cost is computed for the best candidate obtained using a recursive transform for the selected mode.

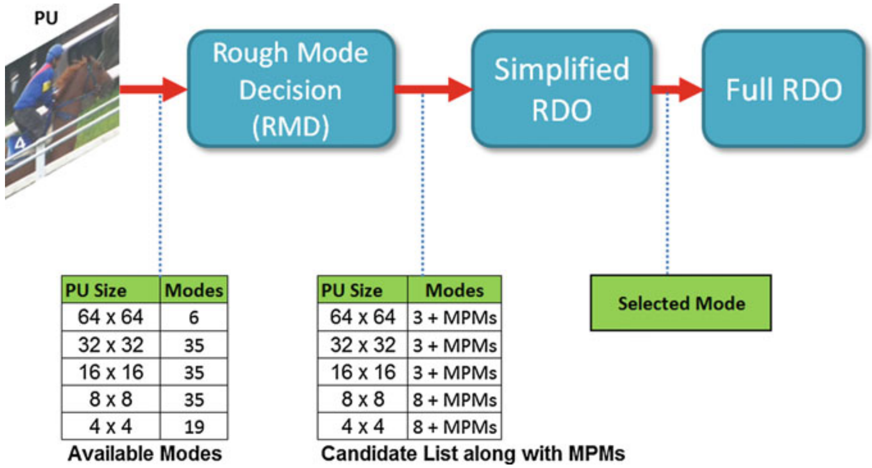

Fig. 4 HEVC mode selection process

The HAD cost for each available mode is computed based on the following formula [9]:

$$J_{HAD} = SATD + \lambda * B \tag{1}$$

where *SATD* represents the Hadamard Sum of Absolute Difference, λ is the Lagrangian constant and *B* specifies the bit cost to be considered in the decision making.

The number of modes available for each PU varies with its size. After comparing the HAD costs of these modes through the RMD process, a list of candidate modes, ordered by increasing cost, is produced.

For the 4 × 4 and 8 × 8 PUs, the RDO process selects the 8 cheapest HAD cost modes. This list is reduced to 3 candidates for larger size PUs [15, 22]. The simplified RDO process compares all the candidates in this list in order to determine the best one for this PU. However, due to this iterative mode search, the complexity is still quite high.

3 Complexity Reduction for Intra Prediction

The high complexity of HEVC for intra prediction can mainly be attributed to two main changes brought in this standard. Firstly, the high number of modes (35 in all) used for each PU compared to only 9 in the previous AVC/H.264 standard requires a larger number of iterations to to be performed prior to identifying the best mode. Secondly, the large LCU of 64 × 64 pixels in the new standard provides numerous combinations of CUs within one LCU, which further adds to the complexity. In the

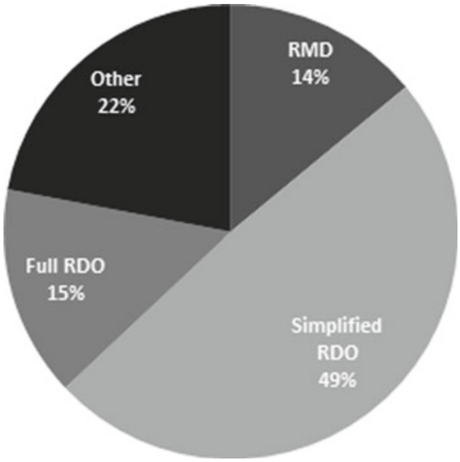

Fig. 5 Time spent during the different stages of HEVC intra prediction

latter case, a considerable time savings could be achieved by identifying the non-split CUs at an early stage of the quad tree traversal to avoid the computation associated with the smaller PU sizes [7, 13, 16, 24].

In this chapter, the focus is on time savings produced through the reduction of the number of modes taking part in the mode-decision process. Figure 5 illustrates the time spent in the different stages during intra prediction. It is observed that 14 % of the encoding time is spent in the RMD process and 49 % in the simplified RDO. In both processes, the selection is made by iterating among a list of modes. The different proposed approaches to reduce the complexity in intra prediction therefore look into techniques to minimize these iterations and yet produce an equivalent mode selection. For each PU, the full RDO process is performed on the selected mode only and this process will not therefore be affected by the mode reduction techniques.

3.1 Commonly Selected Modes

Out of the 35 possible modes from which each PU iterates to find the optimal one, some modes are more commonly selected than others. Figure 6 illustrates the average percentage selection of each of these modes for the list of sequences considered using the 4 Quantization Parameter (QP) values.

Modes 0, 1, 10 and 26 clearly comes out as the most commonly selected modes ranging from 9 % for mode 10 to 26 % for mode 0. These modes have therefore been included by default in the reduced mode approaches proposed for the RMD process. It is to be noted that each sequence may have a different percentage selection for each mode. Figure 7, for example, indicates that more than 55 % of the PUs for the *KristenAndSara* sequence select mode 10. Nevertheless, as illustrated in the figure, the four modes mentioned earlier (0, 1, 10 and 26) remain the four most commonly selected ones.

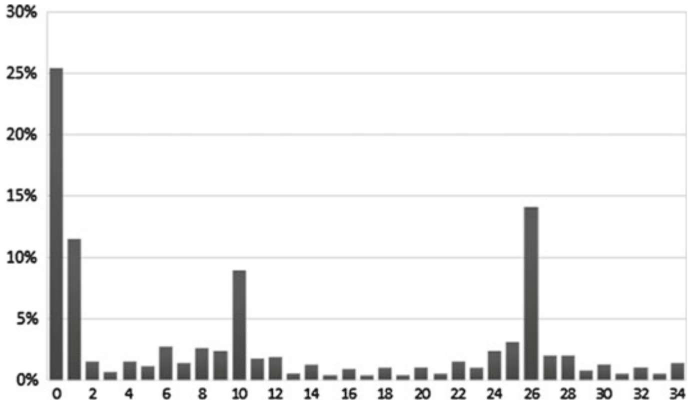

Fig. 6 Percentage occurrence of modes

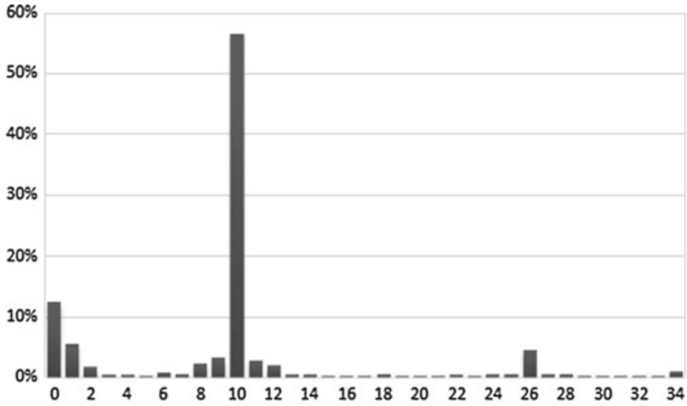

Fig. 7 Percentage occurrence of modes for the sequence *KristenAndSara*

3.2 Experimental Conditions

All experiments in this chapter were conducted using the HEVC Reference Model (HM) version 10 encoder. A list of 21 test sequences with resolution ranging from 416 × 240 (Class E) to 2560 × 1600 (Class A) have been used throughout using the *all intra high-efficiency* encoding configuration with QP values of 22, 27, 32 and 37. The performance of the proposed methods are thus reported in terms of the change in average bitrate, Peak Signal-to-Noise Ratio (PSNR) and total encoding time based on the following formulas:

$$\Delta BitRate(\%) = \frac{BitRate(proposed) - BitRate(HM)}{BitRate(HM)} * 100 \tag{2}$$

$$\Delta PSNR(dB) = PSNR(proposed) - PSNR(HM) \tag{3}$$

$$\Delta Time(\%) = \frac{Time(proposed) - Time(HM)}{Time(HM)} * 100 \tag{4}$$

The Bjøntegaard metrics [1] rate distortion performance is used in the computation of BD-Rate and BD-PSNR.

The use of fewer modes for the RMD process brings significant reduction in complexity. This is further studied in the following section.

4 Reduced Mode Approach for the RMD Process

In this approach, the aim is to provide the RDO process with the same or similar list of candidates, using a faster method. The PLANAR mode (0) and the DC mode (1) are inserted by default in the initial list as they correspond to the highly selected modes. In addition these two modes are non-angular and will not be considered in the reduced mode approach. The reduction in modes therefore applies only to the 33 angular modes. Figure 8 provides an example of the HAD cost for the different modes of an 8 × 8 PU. The graph shows that the HAD cost varies in a relatively continuous waveform, i.e. there is a gradual decrease or increase in HAD cost from one mode to the next and the graph may have a number of minimum values. The graph spans from mode 2 to mode 34. Modes 0 and 1 are also included in the figure. The modes falling below the dotted line are the 8 cheapest cost modes forming part of the candidate list for that PU. The graph clearly illustrates that the initial modes on the graph and those on the right hand side are well above the first 8 cheapest cost modes. The proposed approaches therefore identify these high cost regions and aim towards reducing these unnecessary HAD cost computations which in turn will

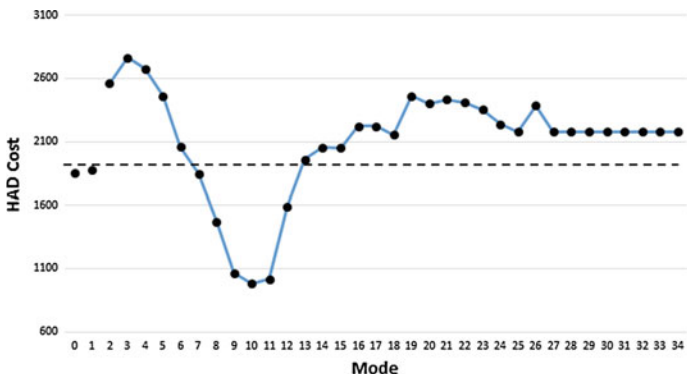

Fig. 8 Plot of HAD cost against mode number for an 8 × 8 PU

lead to a substantial time savings. In the following sub-sections, the approach to reduce the 35 modes and the 19 modes commonly used for the different PUs will be elaborated. The RMD for the 64 × 64 PU has not been optimized as it is already making use of a very reduced number of modes, i.e. 6.

4.1 Reducing the 35 Modes

PU sizes 8 × 8, 16 × 16 and 32 × 32 iterates among the 35 modes during the RMD process in the conventional HEVC. The proposed method to reduce the number of modes is as follows:

1. Initially, a list of modes is constructed starting with mode 2 at intervals of 4. The modes identified are 2, 6, 10, 14, 18, 22, 26, 30 and 34. The HAD cost of these 9 modes are computed to build the list of candidate modes (3 or 8 depending on PU size).
2. At this stage, additional modes which are close to the cheapest modes obtained in the previous step will also be included. The number of cheapest modes that will be considered is denoted by *lval*. The minimum value for *lval* is 1 and the maximum is the number of candidates in the list (3 or 8). The lower the value of *lval*, the fewer will be the number of modes to be considered. However, the improved time savings with a low *lval* may also lead to a performance deterioration.
3. For each of the number of lowest values as set by *lval*, the HAD cost of the 2 left adjacent neighbors and the 2 right adjacent neighbors are identified. For mode 2 and mode 34, which are located at the two extreme ends of the list of angular modes, only the 2 right and 2 left modes respectively are included. The HAD cost of these additional modes, provided they were not processed earlier, are computed and the candidate list is updated accordingly.
4. Finally the HAD cost for mode 0 and for mode 1 are computed and the candidate list is again updated.

The resulting candidate list will therefore contain the same number of modes as in the conventional HEVC although the processing of some of the 35 modes have been skipped. Thus, if only the cheapest cost mode is chosen in step 2 above ($lval = 1$), a maximum of 15 modes will go through the HAD cost computation. Since the number of modes cannot be exactly determined when choosing the value of *lval*, the N_{lval}^{modes} notation will be used, where *modes* denotes the number of available modes for that PU and *lval* indicates the number of low values used in the scheme.

The detailed results of using the N_3^{35} scheme is shown in Table 3. An encoding time gain of 4.1 % is observed with a trivial decrease in PSNR (−0.0009 dB) along with a 0.07 % increase in bitrate. This decrease in encoding time prior to entering the RDO process is quite significant at this stage.

The results obtained with different values of *lval* are indicated in Table 4 for values of *lval* between 1 and 4. For the 16 × 16 and 32 × 32 PUs, only the 3 cheapest modes will be included in the candidate list while for the 8 × 8 PU this number is increased

Table 3 Performance of HM using the N_3^{35} scheme

Resolution	Sequence	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)
2560 × 1600	PeopleOnStreet	−0.0007	−3.8	0.03
	Traffic	−0.0021	−3.8	0.06
1920 × 1080	BQTerrace	−0.0009	−3.5	0.06
	Cactus	−0.0018	−3.7	0.01
	Tennis	−0.0009	−4.5	0.11
1280 × 720	FourPeople	0.0000	−3.9	0.01
	KristenAndSara	−0.0031	−3.5	0.03
	vidyo1	0.0024	−4.3	0.08
	vidyo3	−0.0025	−4.0	0.09
	vidyo4	−0.0004	−4.3	−0.02
832 × 480	BasketballDrill	0.0006	−3.5	0.21
	BQMall	−0.0008	−3.7	0.05
	Flowervase	−0.0032	−4.8	0.12
	Keiba	0.0019	−3.7	−0.01
	PartyScene	−0.0019	−2.8	0.00
	RaceHorses	−0.0015	−3.5	0.03
416 × 240	BasketballPass	−0.0045	−3.3	0.14
	BlowingBubbles	0.0044	−5.2	0.21
	Flowervase	−0.0006	−5.8	0.07
	Keiba	0.0091	−4.7	0.18
	Mobisode2	−0.0134	−5.3	0.04
Average		−0.0009	−4.1	0.07

Table 4 Results of using different schemes for PUs considering 35 modes

Scheme	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)
N_1^{35}	−0.0035	−5.4	0.13
N_2^{35}	−0.0014	−4.3	0.10
N_3^{35}	−0.0009	−4.1	0.07
N_4^{35}	0.0003	−3.8	0.08

to 8 modes. It is observed that an encoding time gain of 5.4 % is obtained by using the N_1^{35} scheme (maximum of 15 modes) along with a slight degradation in quality. Considering that the whole RMD process consumes 14 % of the total encoding time, this 5.4 % overall gain is significant. In the subsequent proposed approaches, the N_1^{35} scheme will be used when time savings is required and the N_3^{35} scheme will be used when quality is of concern rather than the conventional 35 modes PUs.

4.2 Reducing the 19 Modes

PUs of size 4 × 4 pixels make use of only 19 modes in the RMD process according to conventional HEVC due to the high number of such PUs for a single LCU. This smallest PU size is responsible for almost 25 % of the total RMD time. A similar approach to the 35 modes, in Sect. 4.1, is therefore adopted to reduce the number of modes considered. However, in this case, only 1 adjacent neighbor to the left and 1 adjacent neighbor to the right are included. Similarly, for modes 2 and 34, only one neighbor is considered. The same notation, N_{lval}^{modes} will be used. The result of using the N_2^{19} scheme is provided in Table 5 and the summary of results for the different values of *lval* are shown in Table 6.

For the 19 modes PUs, the time savings obtained is relatively small. The reason being that for the N_1^{19} scheme a maximum of 13 modes are already chosen out of 19. It is found, however, that the N_2^{19} scheme provides a good performance of 0.90 % reduction in encoding time together with a small *decrease* in bitrate and a negligible drop in PSNR of 0.0005dB. This scheme will therefore be adopted throughout the

Table 5 Performance of HM using the N_2^{19} scheme

Resolution	Sequence	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)
2560 × 1600	PeopleOnStreet	0.0011	−1.03	0.010
	Traffic	0.0008	−1.03	−0.006
1920 × 1080	BQTerrace	0.0002	−0.72	0.013
	Cactus	−0.0012	−0.72	−0.027
	Tennis	−0.0004	−1.01	0.002
1280 × 720	FourPeople	−0.0012	−0.75	−0.022
	KristenAndSara	−0.0004	−0.45	−0.011
	vidyo1	0.0001	−0.86	0.076
	vidyo3	−0.0021	−0.92	−0.076
	vidyo4	−0.0019	−0.85	−0.086
832 × 480	BasketballDrill	−0.0043	−0.96	−0.060
	BQMall	−0.0019	−0.93	−0.007
	Flowervase	−0.0010	−1.59	0.035
	Keiba	0.0023	−0.93	0.025
	PartyScene	−0.0017	−0.72	−0.022
	RaceHorses	−0.0031	−0.92	−0.059
416 × 240	BasketballPass	−0.0005	−0.62	0.131
	BlowingBubbles	0.0003	−0.29	0.092
	Flowervase	−0.0003	−1.17	0.019
	Keiba	0.0009	−1.18	−0.011
	Mobisode2	0.0029	−1.17	−0.044
Average		−0.0005	−0.90	0.000

Table 6 Results of using different schemes for PU considering 19 modes

Scheme	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)
N_1^{19}	−0.0020	−0.98	0.0099
N_2^{19}	−0.0005	−0.90	−0.0003
N_3^{19}	0.0002	−0.82	0.0247

remaining experiments as it represents almost a 1 % guaranteed gain in encoding time with practically no deterioration in quality at the very start, prior to entering the RDO process.

5 Reducing Modes in the RDO Process

As shown in Fig. 4, HEVC makes use of a candidate list consisting of 3 modes for PU of size 16 × 16 and higher while 8 modes are selected for smaller size PUs. The proportion of time spent in the simplified RDO process for the different PU sizes is illustrated in Fig. 9. This process, which represents 49 % of the time spent for the whole intra prediction, produces the final mode decision for a PU.

A reduction in the number of modes used in this process will therefore lead to a significant time savings. Three approaches are considered:

1. Using fewer modes from the candidate list
2. Using a fraction of the original candidate list based on a threshold
3. Using a fraction of the original candidate list based on a threshold along with a reduced MPM

It can also be noted that the reduced mode approach for the RMD process in Sect. 3 has been integrated in the subsequent experiments. More specifically, the reduced RMD schemes N_1^{35} and N_2^{19} are used as they together provide more than 6 % time savings prior to entering the RDO process.

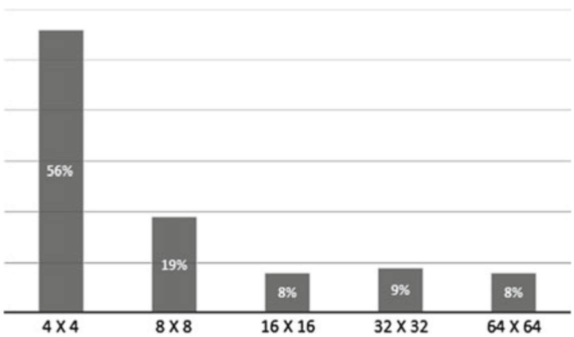

Fig. 9 Percentage time spent in the simplified RDO for each PU size

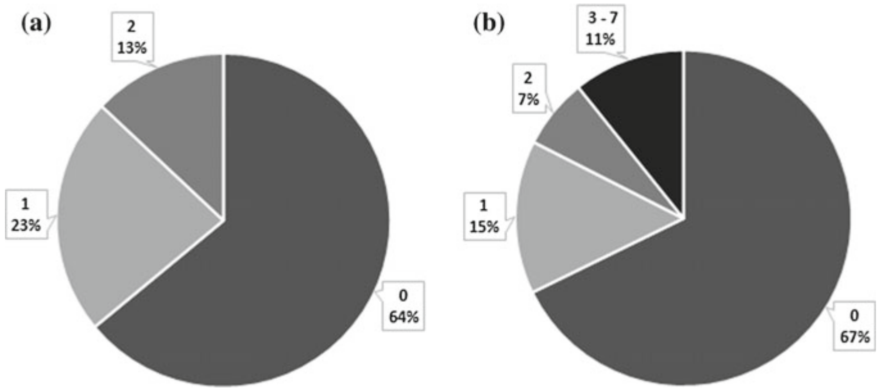

Fig. 10 Distribution of selected RDO modes **a** among the 3 RMD modes for 16 × 16 , 32 × 32 and 64 × 64 PUs and **b** among the 8 RMD modes for 4 × 4 and 8 × 8 PUs

5.1 Relevance of Mode Positioning in the Candidate List

The simplified RDO iterates among the modes from the candidate list along with the MPMs to identify the best mode for a given PU. The candidate list produced by the RMD process is ordered in ascending order of HAD cost starting from the cheapest. Four sequences have been used to collect the position of the selected mode in the original candidate list and the results are shown in Fig. 10a for the 4 × 4 and 8 × 8 (8 modes) and in Fig. 10b for larger size PUs (3 modes).

In this experiment, only the selected modes coming from the RMD process have been taken into account without considering the MPMs. The figures indicate that more than 75 % of all selections come from the first and second positions of the candidate list. In fact, for the small PU sizes, 67 % of them select position 0 (the cheapest cost) from the candidate list while 64 % of the larger PU sizes adopt this selection. Position 0 and position 1 (the first two modes) of the candidate list share together almost 80 % of the selections for smaller PUs and 87 % for the larger ones. It implies that these two positions should imperatively be considered in mode reduction approaches.

5.2 Using Fewer Candidates in the RDO Process

As discussed in Sect. 5.1, the starting modes in the candidate list play a major role in the mode decision process. Three distinct sets of experiments are therefore conducted to evaluate the impact of using 1, 2 and 3 modes only from the candidate list for all PUs in a sequence. The MPMs are then included accordingly. The results of using only the first mode along with the MPMs are thus displayed in Table 7.

Table 7 Performance of HM using only the first RMD mode position along with MPMs

Resolution	Sequence	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)	BD-PSNR	BD-Rate
2560 × 1600	PeopleOnStreet	−0.074	−37.8	0.03	−0.081	1.42
	Traffic	−0.062	−37.9	−0.03	−0.058	1.20
1920 × 1080	BQTerrace	−0.073	−39.8	−0.21	−0.058	1.04
	Cactus	−0.056	−37.9	−0.31	−0.040	1.25
	Tennis	−0.029	−38.9	0.19	−0.038	1.28
1280 × 720	FourPeople	−0.072	−38.3	0.03	−0.089	1.37
	KristenAndSara	−0.030	−40.4	0.01	−0.025	0.42
	vidyo1	−0.060	−38.2	0.15	−0.079	1.47
	vidyo3	−0.045	−38.8	0.39	−0.086	1.37
	vidyo4	−0.049	−38.4	0.03	−0.049	1.18
832 × 480	BasketballDrill	−0.058	−37.2	−0.01	−0.060	1.18
	BQMall	−0.089	−37.8	−0.16	−0.081	1.36
	Flowervase	−0.096	−39.6	0.04	−0.096	1.62
	Keiba	−0.056	−38.6	−0.44	−0.035	0.71
	PartyScene	−0.128	−37.2	−0.40	−0.093	1.34
	RaceHorses	−0.077	−37.1	−0.32	−0.056	0.97
416 × 240	BasketballPass	−0.085	−37.0	0.09	−0.102	1.70
	BlowingBubbles	−0.099	−39.0	−0.32	−0.077	1.38
	Flowervase	−0.086	−40.0	−0.09	−0.077	1.37
	Keiba	−0.047	−40.6	0.02	−0.046	0.74
	Mobisode2	−0.022	−37.9	0.78	−0.058	1.64
Average		−0.066	−38.5	−0.03	−0.066	1.24

An average complexity reduction of 38.5 % is achieved using the first candidate mode only along with the MPMs for the RDO process. This reduction is associated with a 0.03 % *reduction* in bitrate and a drop of 0.066 dB in terms of PSNR. For the list of sequence tested, the average BD-PSNR is only −0.066 dB while the average BD-Rate is 1.24 %. The time savings for each of the sequences show little deviation from the average value indicating the small change in the number of modes coming from the additional MPMs. In the conventional approach, the candidate list may already contain most of the MPMs leaving only the few remaining ones to be added. When only one mode is chosen from the RMD process, most of the MPMs will have to be added to the candidate list and the final list may contain up to 4 modes, including the MPMs. Experiments have thus been conducted by using different number of modes from the candidate list and a summary of their results are depicted in Table 8.

Increasing the number of candidate modes produces lower complexity reduction, as will be expected, and the minor change in bitrate (*decrease*) implies that the compression factor has not been affected compared to conventional HEVC. The BD-PSNR and the BD-Rate also show lower values as additional modes are included.

Table 8 Summary of performance using fewer modes from the candidate list along with MPMs

RMD modes used	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)	BD-PSNR (dB)	BD-Rate (%)
0	−0.066	−38.5	−0.03	−0.066	1.24
1	−0.037	−32.5	−0.08	−0.034	0.65
2	−0.024	−26.9	−0.01	−0.024	0.46

Table 9 Summary of performance using fewer modes in the candidate list with the N_2^{35} and N_2^{19} scheme

RMD modes used	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)	BD-PSNR (dB)	BD-Rate (%)
0	−0.066	−37.8	−0.09	−0.052	1.15
1	−0.036	−32.1	−0.15	−0.028	0.54
2	−0.021	−26.2	−0.07	−0.017	0.33

The results show that considerable complexity reduction can therefore be achieved without increasing the bitrate at varying degrees of PSNR drop by using fewer modes in the candidate list.

The same experiment is conducted, but this time with the N_2^{35} and N_2^{19} reduced RMD schemes. Under these conditions, the number of modes in the RMD process is slightly increased for the 35 modes PUs to achieve a better quality in the output and this is illustrated by the results in Table 9. Although the complexity reduction is only slightly lower (less that 1 % when using N_2^{35}) compared to the N_1^{35} scheme used earlier, the bitrate is further reduced along with marginal change in PSNR. The BD-PSNR and the BD-Rate values are much lower compared to Table 8, confirming a substantial gain in quality although the complexity reduction is lower by less than 1 %.

5.3 Using Variable Number of Candidates in the RDO Process

In Sect. 5.2, a fixed number of modes is utilized from the candidate list produced by the RMD process. Figure 10 pointed out that modes situated at the beginning of the RMD list have a higher probability of being selected and modes in the higher positions are not so frequently used. However, when these higher position modes are selected they contribute significantly to the quality in the compression algorithm. This is readily observed from Tables 8 and 9 where the drop in PSNR becomes less significant compared to conventional HEVC when additional modes are used in the RDO process. A variable number of modes is therefore proposed based on a threshold value. The threshold value is set as a percentage of the difference in HAD

Table 10 Distribution of selected RMD modes with respect to %R for different PU sizes

%R	4 (%)	8 (%)	16 (%)	32 (%)	64 (%)
0–10	70	64	71	62	54
10–20	6	7	3	1	11
20–30	5	5	3	4	6
30–40	4	4	2	4	3
40–50	3	3	1	3	3
50–60	3	4	3	4	0
60–70	3	3	1	1	3
70–80	2	2	2	2	3
80–90	2	2	3	2	3
90–100	3	5	12	17	14

cost between the first mode and the last mode in the candidate list produced by the RMD process. The difference in cost which is termed as the Range, R, is given by:

$$R = HAD_{MaxMode} - HAD_0 \tag{5}$$

where HAD_i represents the HAD cost and i denotes the mode number starting from 0. *MaxMode* will take the value of 7 for the smaller size PUs and 2 for the larger ones. When the threshold value is used, only modes with HAD cost difference of less than %R from the candidate list will be retained for the RDO process along with the MPMs. The distribution of the selected RMD modes with respect to %R for the different PU sizes of the *BasketBallPass* sequence with QP = 22 is displayed in Table 10. The other sequences with different QP values show a similar trend.

Table 10 reveals that most of selected modes have a difference of less than 10 %R. Consequently experiments with threshold values of 5 %R and 10 %R are conducted. The detail results for a threshold value of 5 %R is given in Table 11. A 38.3 % gain in encoding time is achieved with a 0.09 % *decrease* in bitrate along with a drop of only 0.060 dB in PSNR. The bitrate has decreased for most of the sequences and a higher drop in PSNR is observed for the lower resolution ones.

The performances using different values of %R is provided in Table 12. The results show a slight decrease in time savings when the threshold is increased, indicating that more modes are being processed by the RDO. With a threshold of 20 %R, a 36 % complexity reduction is achieved with a drop of only 0.043 dB in PSNR along with a *decrease* of 0.08 % in bitrate. However, although the bitrate remains relatively static (but still with negative values), the quality in terms of PSNR shows much improvement as %R is increased.

The performances using 5 %R as threshold with different values of *lval* for the N^{35}_{lval} scheme are shown in Table 13. The results for the different schemes show only marginal improvement in bitrate with practically no change in quality compared to the N^{35}_{1} scheme.

Table 11 Performance of HM using 5 %R along with MPMs

Resolution	Sequence	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)
2560 × 1600	PeopleOnStreet	−0.064	−37.9	−0.06
	Traffic	−0.055	−37.8	−0.08
1920 × 1080	BQTerrace	−0.067	−39.6	−0.25
	Cactus	−0.049	−37.7	−0.25
	Tennis	−0.026	−38.8	−0.02
1280 × 720	FourPeople	−0.066	−38.0	−0.02
	KristenAndSara	−0.029	−40.4	−0.03
	vidyo1	−0.053	−38.2	0.08
	vidyo3	−0.041	−38.9	0.30
	vidyo4	−0.045	−38.4	−0.07
832 × 480	BasketballDrill	−0.050	−37.1	−0.12
	BQMall	−0.079	−37.6	−0.21
	Flowervase	−0.086	−39.9	−0.06
	Keiba	−0.045	−38.4	−0.32
	PartyScene	−0.119	−36.8	−0.44
	RaceHorses	−0.071	−36.9	−0.39
416 × 240	BasketballPass	−0.066	−37.5	0.11
	BlowingBubbles	−0.091	−36.8	−0.29
	Flowervase	−0.081	−39.4	−0.13
	Keiba	−0.041	−39.9	−0.12
	Mobisode2	0.030	−37.8	0.24
Average		−0.060	−38.3	−0.09

Table 12 Summary of performances using different threshold values

%R	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)
5	−0.060	−38.3	−0.09
10	−0.055	−37.6	−0.12
20	−0.043	−36.0	−0.08

Table 13 Summary of performance using 5 %R as threshold for different values of the N_{lval}^{35} scheme

Reduced RMD scheme	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)
N_1^{35}	−0.060	−38.3	−0.09
N_2^{35}	−0.062	−37.6	−0.18
N_3^{35}	−0.060	−36.8	−0.18

5.4 *Selecting Candidates Based on %R Cost Threshold Along with a Reduced MPM*

In Sects. 5.2 and 5.3, complexity reduction methods have been implemented and performances above 38 % have been achieved with acceptable loss in quality. In [8], it is proposed to skip the MPM coding for the 64 × 64 PUs. In order to increase the time savings further, only the MPMs for PU size of 4 × 4 and 8 × 8 are considered. By skipping the MPMs for large size PUs, the number of RDO iterations will be automatically reduced to increase the time savings. However, PUs of small sizes contribute considerably to the quality of the output and therefore these MPMs are maintained. The detailed results of using a 5 %R along with the proposed reduced MPMs is thus provided in Table 14.

An average reduction of 42.1 % encoding time is thus achieved along with a negligible increase of 0.27 % bitrate and a loss in quality of 0.075 dB. The reduction of complexity shows little variation throughout. It is noted, however, that the sequences *Mobisode2*, *Tennis* and *Video1* have the highest increase in bitrate whereas the two

Table 14 Performance of HM using 5 % R along with reduced MPMs

Resolution	Sequence	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)
2560 × 1600	PeopleOnStreet	−0.081	−42.0	0.20
	Traffic	−0.071	−41.7	0.30
1920 × 1080	BQTerrace	−0.072	−43.6	−0.12
	Cactus	−0.059	−41.6	0.07
	Tennis	−0.048	−42.5	0.92
1280 × 720	FourPeople	−0.082	−42.3	0.31
	KristenAndSara	−0.036	−43.4	0.11
	vidyo1	−0.076	−42.2	0.79
	vidyo3	−0.059	−42.7	0.71
	vidyo4	−0.068	−42.2	0.63
832 × 480	BasketballDrill	−0.067	−41.4	−0.02
	BQMall	−0.086	−42.2	0.01
	Flowervase	−0.116	−43.4	0.33
	Keiba	−0.058	−42.3	0.00
	PartyScene	−0.118	−41.3	−0.37
	RaceHorses	−0.084	−41.0	−0.23
416 × 240	BasketballPass	−0.082	−41.5	0.23
	BlowingBubbles	−0.092	−41.0	−0.03
	Flowervase	−0.128	−42.7	0.43
	Keiba	−0.057	−43.1	0.10
	Mobisode2	−0.029	−40.8	1.36
Average		−0.075	−42.1	0.27

Table 15 Summary of performances using 5 %R and 10 %R as threshold for different values of the N_{lval}^{35} scheme

%R	Reduced RMD scheme	Δ PSNR (dB)	Δ Time (%)	Δ BitRate (%)
5	N_1^{35}, N_2^{19}	−0.075	−42.1	0.27
	N_2^{35}, N_2^{19}	−0.073	−41.5	0.17
	N_3^{35}, N_2^{19}	−0.073	−40.8	0.15
10	N_1^{35}, N_2^{19}	−0.067	−41.6	0.26
	N_2^{35}, N_2^{19}	−0.066	−40.8	0.15
	N_3^{35}, N_2^{19}	−0.067	−40.0	0.11

Flowervase sequences and the *PartyScene* sequence are found to have impacted negatively on the PSNR.

A summary of the performance of HM using different values of %R with different reduced RMD schemes is provided in Table 15. The table shows that a performance of at least 40 % is achieved for all the different combinations of schemes and the different %R used. The drop in quality ranges from −0.066 dB to −0.075 dB while increase in bitrate ranges from 0.11 to 0.27 %.

The performance comparisons of the proposed method for the *Traffic* sequence (2560 × 1600) is shown in Fig. 11. It confirms that the PSNR loss of the proposed algorithm over the full range of QP values is marginal for a complexity reduction of 42.1 %. The experimental results also indicate that the proposed method can efficiently reduce the encoding complexity with little performance loss.

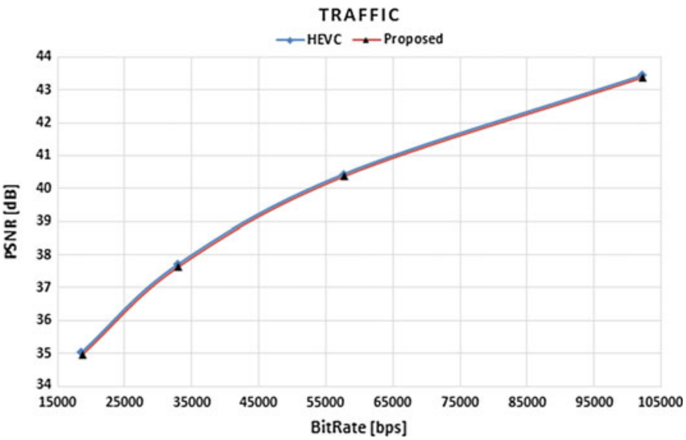

Fig. 11 Coding performance of proposed method based on a 5 %R cost threshold along with a reduced MPM for the *Traffic* sequence

Fig. 12 Output from conventional HM for the sequence Traffic for QP = 22

Fig. 13 Output from proposed method for the sequence Traffic for QP = 22

Figures 12 and 13, relating to the *Traffic* sequence using a QP value of 22, show extracts from the outputs of the conventional HM and the proposed method respectively. The visual quality of the outputs shows no perceptible difference.

The corresponding LCU in each figure is highlighted. It shows that the LCUs in both the conventional HM and the proposed method exhibit similar structure. It is also observed that CUs are of larger sizes in the proposed approach which partly explains the decrease in PSNR observed.

6 Conclusion

In this chapter, various fast mode decision methods for HEVC intra prediction have been proposed. The proposals made were directed towards reducing the number of modes taking part in the RMD and the RDO processes.

The approach used led to a maximum of 5.4 % overall reduction in complexity by reducing modes taking part in the RMD process for the 35 modes PUs and 0.98 %

for the 19 modes PUs. The approaches used for the RDO process include reducing the number of candidates in the RMD list, using a HAD cost threshold based on the difference between the first and last mode in the candidate list and skipping the MPMs for the large size PUs. On the basis of the initial encoding time reduction during the RMD and of techniques to reduce the number of iterations in the RDO process, various approaches have been proposed where complexity reduction of more than 38 % are achieved with an acceptable loss in bitrate and PSNR.

By using only the first RMD mode, a 38.5 % reduction in encoding time is achieved with a PSNR drop of 0.066 dB along with a 0.03 % *decrease* in bitrate. Using a 5 %R HAD cost threshold, a 38.3 % time savings is achieved with a drop in PSNR of only 0.060 dB and a 0.09 % *decrease* in bitrate. The maximum complexity reduction is reached when a 5 %R HAD cost threshold is used together with a reduced MPM. The performance indicates a 42.1 % time savings along with a drop of 0.075 dB in terms of PSNR and a negligible increase of 0.27 % in bitrate.

References

1. Bjøntegaard, G.: Calculation of Average PSNR Differences between RDCurves. Document VCEG-M33, VCEG (2001)
2. Bossen, F., Bross, B., Suhring, K., Flynn, D.: HEVC complexity and implementation analysis. IEEE Trans. Circuits Syst. Video Technol. **22**(12), 1685–1696 (2012)
3. Chen, G., Liu, Z., Ikenaga, T., Wang, D.: Fast HEVC intra mode decision using matching edge detector and kernel density estimation alike histogram generation. In: 2013 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 53–56 (2013)
4. Chen, G., Pei, Z., Sun, L., Liu, Z., Ikenaga, T.: Fast intra prediction for HEVC based on pixel gradient statistics and mode refinement. In: 2013 IEEE China Summit International Conference on Signal and Information Processing (ChinaSIP), pp. 514–517 (2013)
5. da Silva, T.L., Agostini, L.V., da Silva Cruz, L.A.: Speeding up HEVC intra coding based on tree depth inter-levels correlation structure. In: 2013 Proceedings of the 21st European Signal Processing Conference (EUSIPCO), pp. 1–5 (2013)
6. da Silva, T.L., da Silva Cruz, L.A., Agostini, L.V.: Fast HEVC intra mode decision based on dominant edge evaluation and tree structure dependencies. In: 2012 19th IEEE International Conference on Electronics, Circuits and Systems (ICECS), pp. 568–571 (2012)
7. Goswami, K., Kim, B.-G., Jun, D., Jung, S.-H., Choi, J.S.: Early coding unit-splitting termination algorithm for high efficiency video coding (HEVC). ETRI J. **36**(3), 407–417 (2014)
8. ISO/IEC JTC 1 SC29 WG11. Prediction modes and mode coding for large size Intra block. Document JCTVC-IO227_v9, JCTVC (2012)
9. ISO/IEC JTC 1 SC29 WG11. High Efficiency Video Coding (HEVC) Test Model 10 (HM10) Encoder Description. Document JCTVC-L1002_v3, JCTVC (2013)
10. ISO/IEC JTC 1 SC29 WG11. High Efficiency Video Coding (HEVC) Text Specification Draft 10. Document JCTVC-L1003_v9, JCTVC (2013)
11. ITU-T. High efficiency video coding. Recommendation ITU-T H.265, ITU (2013)
12. Jiang, W., Ma, H., Chen, Y.:. Gradient based fast mode decision algorithm for intra prediction in HEVC. In: 2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet), pp. 1836–1840 (2012)
13. Kim, J., Choe, Y., Kim, Y.G.: Fast Coding Unit size decision algorithm for intra coding in HEVC. In: 2013 IEEE International Conference on Consumer Electronics (ICCE), pp. 637–638 (2013)

14. Kumar, V., Govindaraju, H., Quaid, M., Eapen, J.: Fast intra mode decision based on block orientation in high efficiency video codec (HEVC). In: 2014 International Symposium on Computer, Consumer and Control (IS3C), pp. 506–511 (2014)
15. Lainema, J., Han, W.-J.: Intra-picture prediction in HEVC. In: Sze, V., Budagavi, M., Sullivan, G.J., (eds.) High Efficiency Video Coding (HEVC), Integrated Circuits and Systems, pp. 91–112. Springer International Publishing, Berlin (2014)
16. Lin, Y.C., Lai, J.C.: Edge density early termination algorithm for HEVC coding tree block. In: 2014 International Symposium on Computer, Consumer and Control (IS3C), pp. 39–42 (2014)
17. McCann, K., Han, W.-J., Kim, I.-K.: Samsungs Response to the Call for Proposals on Video Compression Technology. Document JCTVC-A124, JCTVC (2010)
18. Piao, Y., Min, J., Chen, J.: Encoder Improvement of Unified Intra Prediction. Document JCTVC-C207, JCTVC (2010)
19. Saurty, K., Catherine, P.C., Soyjaudah, K.M.S.: A modified rough mode decision process for fast high efficiency video coding (HEVC) intra prediction. In: 2015 Fifth International Conference on Digital Information and Communication Technology and its Applications (DICTAP), pp. 143–148 (2015)
20. Shi, Y., Au, O.C., Zhang, H., Zhang, X., Jia, L., Dai, W., Zhu, W.: Local saliency detection based fast mode decision for HEVC intra coding. In: 2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSP), pp. 429–433 (2013)
21. Sullivan, G.J., Ohm, J., Han, W.-J., Wiegand, T.: Overview of the high efficiency video coding (HEVC) standard. IEEE Trans. Circuits Syst. Video Technol. **22**(12), 1649–1668 (2012)
22. Sun, H., Zhou, D., Goto, S.: A low-complexity HEVC intra prediction algorithm based on level and mode filtering. In: 2012 IEEE International Conference on Multimedia and Expo (ICME), pp. 1085–1090 (2012)
23. Wiegand, T., Sullivan, G.J., Bjontegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. IEEE Trans. Circuits Syst. Video Technol. **13**(7), 560–576 (2003)
24. Zhang, H., Ma, Z.: Early termination schemes for fast intra mode decision in high efficiency video coding. In: 2013 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 45–48 (2013)