Accepted in Data Compression Conference (DCC 2015), 2015

Fast HEVC Intra Mode Decision Based on Edge Detection and SATD Costs Classification

Mohammadreza Jamali[1], Stéphane Coulombe[1], François Caron[2]
[1]*École de technologie supérieure, Université du Québec, Montréal, Canada*
[2]*Vantrix Corporation, Montréal, Canada*
mohammadreza.jamali.1@ens.etsmtl.ca, stephane.coulombe@etsmtl.ca,
francois.caron@vantrix.com

Abstract: The recent High Efficiency Video Coding (HEVC) standard was designed to achieve significantly improved compression performance compared to the widely used H.264/AVC standard. This achievement was motivated by the ever-increasing popularity of high-definition video applications and the emergence of ultra-HD. Unfortunately, this comes at the expense of a significant increase in computational complexity for both inter and intra coding. To alleviate this problem, in this paper, we propose a fast intra mode decision method based on improved edge detection, consideration of most relevant modes from neighboring blocks, and classification of SATD costs permitting the elimination of several candidate modes prior to rate distortion optimization (RDO). Experimental results show that the proposed method provides time reduction up to 39.2% and an average 35.6% with negligible quality loss as compared to the HEVC reference implementation HM 15.0.

1. Introduction

The new video coding standard, H.265/HEVC [1, 2], can reduce the bit rate by half relative to H.264/AVC for the same visual quality. HEVC, like previous video coding standards, is a hybrid video coding scheme that uses improved features in all encoding modules to enhance the coding efficiency. In intra coding, compared to H.264's 9 modes, HEVC employs 33 directional modes for angular prediction in addition to DC and planar modes [3]. This number of modes is now adequate to more precisely predict the directional patterns in high-resolution videos when large block sizes are used. However, this, in addition to the transform tree optimization, has rendered the HEVC intra coding operation much more complex, as compared to previous video codecs.

In mode decision, many computations are needed to select the best intra mode by performing the highly demanding rate distortion optimization (RDO) process. To address this problem, during the development of the standard and after its finalization, several research studies have proposed various approaches for fast intra mode decision. First, in Piao et al. [4], the authors proposed rough mode decision (RMD) as a step prior to RDO to exclude the majority of modes from the highly expensive RDO. Their approach reduces the encoding time by about 50%, and has been used in the HEVC reference implementation since its early versions. Down the line, Zhao et al. [5] further reduced the number of candidates by introducing the concept of most probable modes (MPM). They achieved time savings of up to 28% using HM 0.9 with almost equivalent rate-distortion performance. The HEVC reference implementation HM 4.0 and the versions following it include Zhao's approach. In [6], Jiang et al. proposed a fast mode decision approach based on the gradient information of pixels and obtained another 20% time reduction in comparison with HM 4.0. Chen et al. [7] improved the gradient method by introducing a mode decision approach based on pixel gradient statistics and mode refinement. They applied this to HM 8.0 and achieved a 28% time reduction. Recently, Zhang et al. [8] proposed a method for fast intra coding and achieved significant time reduction by using the Hadamard cost-based progressive rough mode search and early coding unit split termination. Since their method involves more than one module and our method includes only mode decision, the extra modules can be added to our method to achieve higher time reduction. The same goes for [9] by Zhao et al., who proposed to reduce time consumption by excluding some coding unit and transform unit depths.

In this paper, we present several innovations contributing to further reductions of time complexity in intra coding while maintaining quality. First, we propose an enhanced gradient-based edge detector. For any detected edge, we assign weights to three adjacent modes to improve the accuracy of the method compared to previous approaches. Then, the most relevant modes are used to improve previous schemes which employ neighboring blocks for mode decision. We then perform a binary classification, based on the sum of absolute transformed differences (SATD) costs, and eliminate less promising candidates. This step contributes to significantly reduce the number of candidates subject to the complex RDO process, and thus provides considerable time reduction. Finally, to achieve further time savings, we bypass the RDO process completely when the mode with the lowest SATD meets certain specific conditions.

This paper is organized as follows. Section 2 provides an overview of the intra coding process in the HEVC standard, including the block partitioning structure, intra prediction, and mode decision. Section 3 presents the proposed method for fast intra coding. Experimental results are shown in section 4, and section 5 concludes the paper.

2. HEVC Intra Mode Coding

The intra coding architecture in H.265/HEVC is basically similar to the previous H.264/AVC standard. Its improvements include more flexible ways of splitting a frame into blocks for performing predictions and transforms, an increased number of modes, a higher range of coding block sizes and adaptive smoothing of the reference samples among others [3]. While the main coding block in H.264/AVC is a macroblock of size 16×16, HEVC uses a more flexible quadtree structure based on a block called the Coding Tree Block (CTB). A CTB is split into Coding Blocks (CBs), Prediction Blocks (PBs) and Transform Blocks (TBs) to perform prediction and transform. This splitting is more flexible, and is especially useful for higher resolution videos.

In HEVC, 35 intra modes can be used, including angular, DC and planar modes. DC mode is used for predicting the homogeneous regions and planar mode is used to produce smooth sample surfaces. Fig. 1 shows HEVC luma intra prediction modes. In the first versions of the standard the number of modes was dependent on the size of PB and only a subset of modes could be selected. Now, however, for all sizes of PB, all 35 modes are tested. The decision for coding a block as intra is made at the CB level, but the intra mode is selected for each PB, and it is possible for PBs in the same CB to have different intra modes. After the intra mode is selected, the prediction is done for TBs inside the

PB. This means that in determining the prediction signal, the spatially neighboring TBs are used [2]. For a TB with size N × N, there are $4N + 1$ samples for prediction from the above, above-right, above-left, left and below-left TBs. Samples from the below-left TB are not always available, and can only be used when they have been processed and decoded beforehand. As an example a prediction by mode 2 is shown in Fig. 1.

HEVC uses the same process for all supported block sizes, which is highly desirable because of the variety of TB sizes and directional modes present. The intra mode decision consists of two major processes: RMD, which selects modes with the lowest SATD costs based on the Hadamard transform, to reduce the number of modes for rate-distortion step, and RDO, which chooses the best mode with the lowest rate-distortion cost among the selected modes by RMD. The rate-distortion cost is defined as:

$$J_{RDO} = D_{SSE} + \lambda_{RDO} \times B$$

where D_{SSE} is the distortion and B is the number of bits needed for coding a PB by a specific mode and λ_{RDO} is a factor related to the quantization parameter (QP).

3. Proposed Fast Intra Mode Decision Algorithm

In this section, we present our fast intra mode decision algorithm for HEVC to reduce the encoding complexity in intra mode selection. Fig. 2 shows the overall block diagram of the system. First, using the Sobel operator, we determine the dominant edges from which we identify the most powerful modes by considering three adjacent modes for each detected edge. In the next two steps, we add DC and planar modes, but also some other relevant modes from five neighboring blocks, to enrich the selected modes. Following that, we order the list of modes obtained based on their SATD costs, exclude some costly ones, and apply a binary classification on the remaining modes to separate the promising ones from the others. Finally, we apply RDO dodging and rate-distortion optimization to find the best intra mode. The classification and RDO dodging are proposed as novel techniques in this paper, and the other steps have been improved in comparison to previous research studies. Subsequent sections explain these steps in detail.

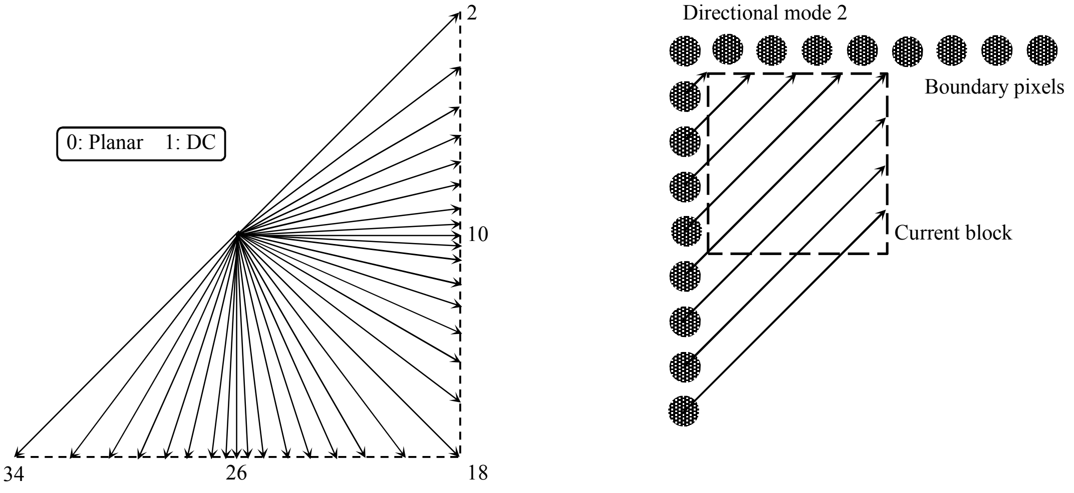

Figure 1: HEVC intra modes

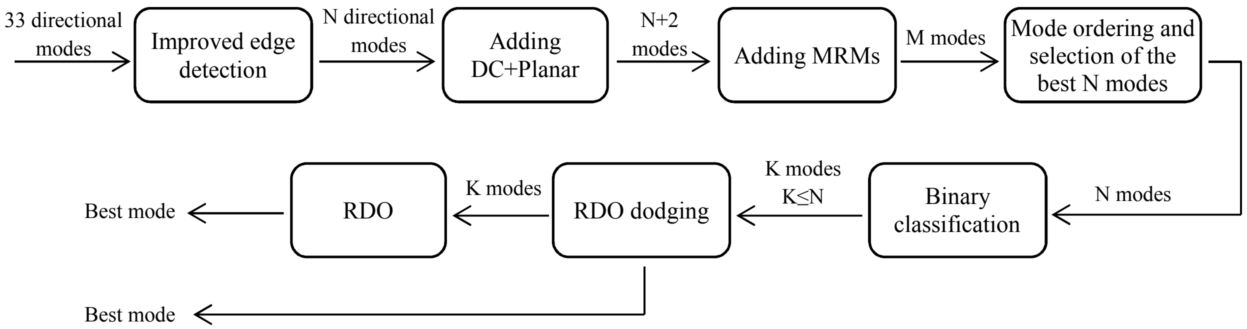

Figure 2: Block diagram of the proposed algorithm

3.1 Improved edge detection based on Sobel operator

In the current reference implementation of HEVC, HM 15.0, the coding process for PBs starts at the CTB level. From there, the process tests all combinations of PBs and all modes for each of them at different depths. Finally, the best depth, the best PB sizes and the best modes for each PB are selected for the CTB. With this in mind, we determine gradients for each pixel at the CTB level. This allows us to use the edge information for each pixel at any depth and prevent a repetition of the calculations for each PB. Using the gradient, we are able to determine the directions with maximum variation of pixel values. The picture edges are perpendicular to these directions, and they show the dominant angular modes for intra prediction. To compute the gradient, we use the Sobel operator with 3×3 convolution masks, and based on these masks, the two components of the gradient are calculated as follows:

$$\vec{G} = G_x\vec{j} + G_y\vec{i}$$

$$G_x = p_{i-1,j+1} + 2 \times p_{i,j+1} + p_{i+1,j+1} - p_{i-1,j-1} - 2 \times p_{i,j-1} - p_{i+1,j-1}$$

$$G_y = p_{i+1,j-1} + 2 \times p_{i+1,j} + p_{i+1,j+1} - p_{i-1,j-1} - 2 \times p_{i-1,j} - p_{i-1,j+1}$$

$$|\vec{G}| = \sqrt{G_x^2 + G_y^2} \quad , \quad Ang(\vec{G}) = atan\left(\frac{G_y}{G_x}\right)$$

We use $|G_x| + |G_y|$ as an estimation of the amplitude and G_y/G_x instead of $Ang(\vec{G})$ because the calculations of *square root* and *atan* are resource-intensive. This data is passed to the lower depths to be used in each PB in selecting the best angular modes. At the PB level, based on the gradient information, a main mode is associated for each pixel. The main mode is defined as the closest mode corresponding to the edge. This mode is achieved by comparing G_y/G_x to predefined limits that are pre-calculated based on the specified angles for intra modes in the HEVC standard. The high and low limits for each mode are shown in Table 1. To be more accurate in finding modes that provide the best prediction, two adjacent modes are considered, in addition to the main mode. This avoids considering only the main mode as the edge is rarely perfectly aligned with it. Therefore, we also give weights to modes adjacent to the main mode based on the direction of the detected edge, as is illustrated in Fig. 3. The weights that are given to each of these modes are calculated as follows.

$$mainModeWeight = \,|\,G_x| + |G_y|$$

$$modeWFP = (atan(highLimit)\ - atan\left(\frac{G_y}{G_x}\right))/(atan(highLimit)\ - atan(lowLimit))$$

$$modeWeightFactor = (highLimit\ - \frac{G_y}{G_x})/(highLimit\ - lowLimit)$$

$$adjacentMode1Weight = (1 - modeWeightFactor) \times (|\,G_x| + |G_y|)$$

$$adjacentMode2Weight = modeWeightFactor \times \left(|\,G_x| + |G_y|\right)$$

In the above formulas, *modeWFP* is the perfect form of the mode weight factor and gives weights of 0.5 for two adjacent modes if the edge sits exactly on a mode. However, as previously mentioned, computing the $atan$ function is avoided, and an approximated formula is used (*modeWeightFactor*). Also, the horizontal mode (mode 10) is considered as a special case, and the weight factor for this mode is as follows:

$$modeWeightFactor = 0.5 \times (1 + 40.73548/\left|\frac{G_y}{G_x}\right|\,)$$

This is because for this mode, there is infinity for the low limit or high limit and the normal formula cannot be applied in this situation. Following these computations, we accumulate the weights of the modes by checking all pixels of the PB, while each pixel increases the weight of three modes. Finally, we obtain a histogram that shows the accumulated weights of 33 directional modes for the whole PB. From this histogram, we consider the *N* most powerful modes as the best candidates selected by edge detection and go to the next steps to add some other modes to enrich this set of *N* modes. The parameter *N* is determined based on the block size (see section 4).

Table 1: High and low limits of G_y/G_x for angular modes

Mode	lowLimit	highLimit	Mode	lowLimit	highLimit
2	-1.15928	-1	18	0.86261	1.15928
3	-1.53711	-1.15928	19	0.65057	0.86261
4	-1.98666	-1.53711	20	0.50336	0.65057
5	-2.59240	-1.98666	21	0.38574	0.50336
6	-3.61354	-2.59240	22	0.27674	0.38574
7	-5.76314	-3.61354	23	0.17352	0.27674
8	-11.61240	-5.76314	24	0.08611	0.17352
9	-40.73548	-11.61240	25	0.02455	0.08611
10	-∞	-40.73548	26	-0.02455	0.02455
10	40.73548	∞	27	-0.08611	-0.02455
11	11.61240	40.73548	28	-0.17352	-0.08611
12	5.76314	11.61240	29	-0.27674	-0.17352
13	3.61354	5.76314	30	-0.38574	-0.27674
14	2.59240	3.61354	31	-0.50336	-0.38574
15	1.98666	2.59240	32	-0.65057	-0.50336
16	1.53711	1.98666	33	-0.86261	-0.65057
17	1.15928	1.53711	34	-1	-0.86261

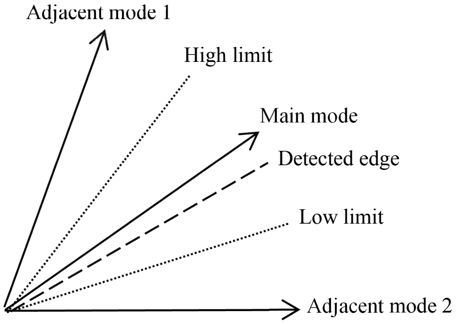

Figure 3: Detected edge and three related modes

3.2 Adding DC, planar and most relevant modes (MRMs)

Since DC and planar are very probable modes in intra prediction [10], and because the edge detection algorithms can only check the angular modes, we add these two modes to our list of candidate modes from the previous step. Then, to exploit the spatial correlation among blocks, we check the best modes of the five neighboring blocks and add them to the list if they are relevant; i.e., if based on their direction, they are likely to be the best mode of the current block. For example, if the mode of the above block is vertical or near-vertical, then it is added to the list. We do the same for the left block if its mode is horizontal or near-horizontal. Similarly, the above-right, above-left and below-left blocks are processed in the same manner. For each of these neighboring blocks, $2n + 1$ modes are defined, which we call most relevant modes (MRMs), and if a neighboring block's intra mode corresponds to one of these modes, then it is considered for the next step. Fig. 4 shows the concept of the most relevant modes of the neighboring blocks. For example, for the above-left block with $n = 2$ the MRMs are 16, 17, 18, 19 and 20, and if that block selects one of these modes, then the mode is added to the list of candidates for the current block. In this step, in addition to DC and planar, a minimum of zero and a maximum of five modes are added to our list of selected modes, and so there are M modes, where $N + 2 \leq M \leq N + 7$.

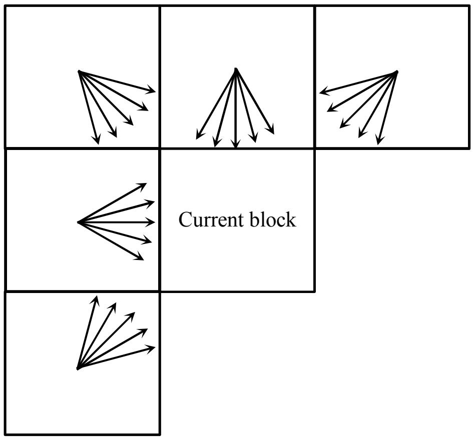

Figure 4: Most relevant modes of the neighboring blocks

3.3 Mode ordering, binary classification and RDO dodging

After adding MRMs, we order the selected modes from the previous step based on the SATD costs, going from the lowest cost to the highest cost, and select the first N modes. The modes are then classified into two classes: powerful contenders with lower costs and weak ones with higher costs. We keep the powerful modes and pass them on to the next step and exclude the weak modes from further processing. This classification could be based on different criteria, but we have found that if a dominant gap exists among the SATD costs, we can efficiently remove the modes with higher costs from the very time-consuming RDO process without affecting the rate-distortion performance. In other words, it is wasteful to test candidates whose costs are much higher than others'. For example, if two candidates have very low costs while the others have significantly higher costs, it is often a waste of time to evaluate all candidates by the computationally expensive RDO, as only the two low-cost candidates usually win. Fig. 5 shows a detected gap among SATD costs where C_{min} is the lowest cost and C_{max} is the highest cost. Finding such a gap is very helpful, and can significantly decrease the time consumption for mode decision in intra coding. It reduces the number of modes to K, where $K \leq N$. $K = N$ occurs when no gap is found. This gap could be a fixed one for all situations or could be adaptive based on the block size, quantization parameter, or absolute minimum and maximum cost values or any other criterion. In this paper, we consider adaptivity based on block size, and the difference between the maximum and minimum cost values. Other parameters could be investigated in future works. The gap is defined as follows:

$$Gap = \alpha \times (C_{max} - C_{min})$$

In this formula, $\alpha \leq 1$ is an empirical parameter, and is adjusted based on the block size (see section 4). Before entering the RDO, we apply another step which we call RDO dodging, which again exploits the modes from five neighboring blocks to bypass the RDO process under certain conditions. The idea is that if the mode with the lowest SATD cost is one of the most relevant modes, then it is highly likely to be the best intra mode for the current block. For instance, if the lowest cost candidate for the current block is the vertical mode and the best mode of the above block is also vertical, we choose this mode as the final decision, and RDO is omitted. To create the set of most relevant modes at this step, we consider $2m + 1$ modes from neighboring blocks similar to section 3.2. If the candidate with the lowest cost after ordering and classification is one of these relevant modes, it is selected as the best mode; otherwise, all K candidates are tested by RDO and the candidate with lowest RDO cost is selected as the final mode.

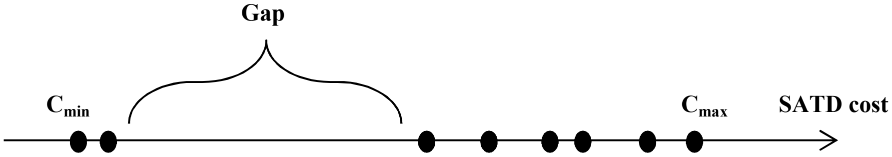

Figure 5: A gap among SATD costs

4. Experimental Results

The proposed algorithm was implemented in the HEVC test model HM 15.0. The implementation platform was an Intel® i7-3770 CPU-3.40, 12 GB of RAM, running Windows 7. We used the 100 first frames of the recommended sequences in [11] to implement our proposed algorithm. HM was configured in *All-Intra* mode, and run for quantization parameters (QP) 22, 27, 32 and 37. The parameters of the algorithm were: $N = 8$ and $\alpha = 1/4$ for block sizes 4×4 and 8×8 and $N = 3$ and $\alpha = 2/3$ for block sizes 16×16, 32×32 and 64×64. Also, n and m were set to 3 and 1 for selecting the most relevant modes. Table 2 shows the results of these experiments in terms of BD-Rate, BD-PSNR_Y [12] and time reduction, in comparison with HM 15.0. Using Bjontegaard metrics, the table shows the average differences in rate-distortion performance. The time reduction is calculated for each QP, and the average over all QPs is presented in the table. According to the experimental results, we achieve an average 35.6% time reduction over all video sequences, and up to 39.2% in comparison with the anchor implementation. The main contributors to this time reduction are a decrease in the number of modes for SATD calculations in the RMD step and a decrease in the number of modes in the extremely time-consuming RDO process. Compared to other works that focus on intra mode decision, our method achieves higher time reductions, as [6] and [7] provide about 20% and 28%, respectively (with 0.04 dB and 0.038 dB loss in BD-PSNR). Moreover, the two main components of our method, binary classification and RDO dodging, are very easy to implement. The price we pay for this time reduction is, on average, a 0.059 dB loss in BD-PSNR_Y and a 1.07% increment in the BD-Rate, which only slightly affects the rate-distortion performance. To justify this, Fig. 6 shows the RD curves of the proposed algorithm versus HM 15.0 for the RaceHorses sequence. From this figure, we can see that both implementations have almost the same rate-distortion performance.

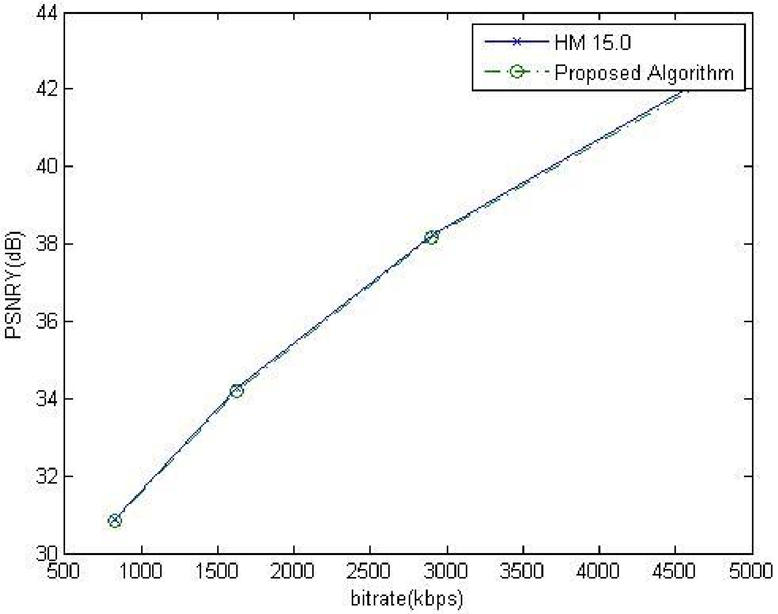

Figure 6: RD curve of the proposed method and HM 15.0 for the RaceHorses sequence

Table 2: Experimental results for recommended video sequences

Class	Video Sequences	ΔT (%)	BD-Rate (%)	BD-PSNR$_Y$ (dB)
A	Traffic	-35.4	0.95	-0.051
	PeopleOnStreet	-34.1	1	-0.057
	NebutaFestival	-34.1	0.53	-0.039
	SteamLocomotiveTrain	-37.8	0.48	-0.025
B	Cactus	-36.1	1.34	-0.05
	Kimono	***-39.2***	0.79	-0.028
	ParkScene	-37.5	0.87	-0.039
	BasketballDrive	-38.4	2.17	-0.059
	BQTerrace	-35.2	0.79	-0.048
C	BQMall	-34.3	1.15	-0.068
	PartyScene	-32.8	1.18	-0.092
	RaceHorsesC	-34.7	0.72	-0.047
	BasketballDrill	-33.1	0.8	-0.039
D	RaceHorses	-34.1	0.99	-0.065
	BasketballPass	-36	1.45	-0.085
	BlowingBubbles	-33.7	1.01	-0.06
	BQSquare	-32.7	1.38	-0.123
E	Vidyo1	-36.8	1.31	-0.066
	Vidyo3	-37.4	1.23	-0.069
	Vidyo4	-37.7	1.33	-0.061
Average		***-35.6***	***1.07***	***-0.059***

5. Conclusion

In this paper, we have presented a fast intra mode decision algorithm for the new video coding standard HEVC. This algorithm improves the edge detection approaches by attributing three modes for each detected edge, and enhances the schemes that exploit the spatial correlation among neighboring blocks by introducing the concept of most relevant modes. Further, two novel techniques, binary classification and RDO dodging, are presented, that significantly reduce the number of calculations for RDO process. Experimental results show that in comparison with the reference implementation of HEVC, HM 15.0, our proposed approach, achieves a 35.6% time reduction on average, with a very small penalty in coding efficiency (0.059 dB loss and 1.07% rate increment). Moreover, in all its steps, the algorithm is based on parameters that can be adjusted based on the desired compromise between complexity reduction and coding efficiency. These make the proposed method attractive for various applications.

Acknowledgment

This work was funded by Vantrix Corporation and by the Natural Sciences and Engineering Research Council of Canada under the Collaborative Research and Development Program (NSERC-CRD 428942-11).

References

[1] B. Bross, W. J. Han, J. R. Ohm, G. J. Sullivan, Y. K. Wang and T. Wiegand, "High Efficiency Video Coding (HEVC) Text Specification Draft 10", *document JCTVC-L1003, ITU-T/ISO/IEC Joint Collaborative Team on Video Coding (JCT-VC)*, Jan. 2013.

[2] G. J. Sullivan, J. R. Ohm, W. J. Han and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 22, No. 12, Dec. 2012.

[3] J. Lainema, F. Bossen, W. J. Han, J. Min and K. Ugur, "Intra Coding of the HEVC Standard", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 22, No. 12, Dec. 2012.

[4] Y. Piao, J. H. Min and J. Chen, "Encoder Improvement of Unified Intraprediction", *JCTVC-C207,* Guangzhou, China, Oct. 2010.

[5] L. Zhao, L. Zhang, X. Zhao, W. Ma, D. Zhao and W. Gao, "Further Encoder Improvement of Intra Mode Decision", *JCTVC-D283,* Daegu, Korea, Jan. 2011.

[6] W. Jiang, H. Ma, Y. Chen, "Gradient Based Fast Mode Decision Algorithm for Intra Prediction in HEVC", *CECNet,* Apr. 2012.

[7] G. Chen, Z. Pei, L. Sun, Z. Liu and T. Ikenaga, "Fast Intra Prediction for HEVC based on Pixel Gradient Statistics and Mode Refinement", *ChinaSIP,* Beijing, July 2013.

[8] H. Zhang and Z. Ma, "Fast Intra Mode Decision for High Efficiency Video Coding (HEVC)", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 24, No. 4, Apr. 2014.

[9] L. Zhao, X. Fan, S. Ma, D. Zhao, "Fast Intra Encoding Algorithm for High Efficiency Video Coding", *Elsevier Signal Processing: Image Communication,* Vol. 29, Issue 9, Oct. 2014.

[10] S. Yan, L. Hong, W. He and Q. Wang, "Group-Based Fast Mode Decision Algorithm for Intra Prediction in HEVC", *International Conference on Signal Image Technology and Internet Based Systems*, Nov. 2012.

[11] F. Bossen, "Common Test Conditions and Software Reference Configurations", *JCTVC-L1100,* 12th Meeting, Geneva, Jan. 2013.

[12] G. Bjontegaard, "Calculation of Average PSNR Differences between R-D Curves", VCEG-M33, Apr. 2001.