Optimized Intra Mode Decision for High Efficiency Video Coding

Anis BenHajyoussef(✉) and Tahar Ezzedine

Ecole Nationale D'Ingénieurs de Tunis, Université de Tunis El Manar,
LR-99-ES21 Sys'Com, 1002 Tunis, Tunisie
anis.hajyoussef@gmail.com, tahar.ezzedine@enit.rnu.tn

Abstract. In order to reach higher coding efficiency, the design of the newest video compression standard - High Efficiency Video Coding (HEVC) - is relying on many improved coding tools and sophisticated techniques. Such a complexity leads to the vital need of video encoders for fast algorithms to overcome the real-time encoding constraint and memory limits. In this context, we propose a gradient based pre-processing stage that will help decreasing the complexity of the encoder and will speed up the Intra mode decision. For that purpose, we investigate the potential of a Prewitt operator instead of the famous Sobel operator used to generate the gradient. In addition and in order to enhance the gradient potential of Intra mode detection without adding a significant complexity, we propose an optimized Intra mode selection through a neighbor mode extension as well as an adapted cost function to take into account the appearing number of modes and the gradient magnitudes. The obtained results demonstrate that we can reduce the encoding time for All Intra configuration by 31.9% with a loss in BD-rate of only 1.1%.

Keywords: HEVC · Intra mode decision · Pre-processing · Image gradient · Sobel · Prewitt · Differential operator

1 Introduction

The emergence of the previous standard, the H.264/AVC, has contributed to an expansion of the video applications. Such an expansion has led to an increasing need for better video quality and higher compression especially with the applications dealing with high and Ultra-High resolutions.

In 2013, HEVC has been developed by the Joint Collaborative Team on Video Coding (JCT-VC), a team of experts from the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG). The new standard keeps the same high-level design as its predecessor but relies on many improved coding tools and techniques that offer higher coding efficiency but at the cost of more encoding complexity. The block structure is one of the most important new features that contributes to this complexity and directly affects all the other features.

V. Murino and E. Puppo (Eds.): ICIAP 2015, Part II, LNCS 9280, pp. 309–319, 2015.
DOI: 10.1007/978-3-319-23234-8_29

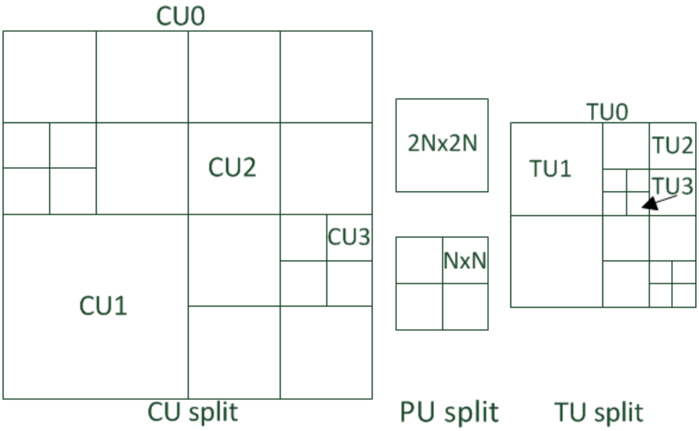

Fig. 1. Example of CUk, PUk and TUk recursive split structures for intra case with k as depth index

In, fact, HEVC relies on a coding tree block (CTB) structure. Unlike the AVC Macrobloc with the size of 16x16, the large coding unit (LCU) defined in HEVC, allows to use block sizes of 8x8 up to 64x64. The LCU can then be partitioned into coding units (CUs) with a quad-tree structure.

At the prediction stage, each CU can be split into one or more prediction units (PUs) [1][2]. Moreover at the transform stage, each CU can be split into one or more recursive transform units (TUs). Figure 1 illustrates a description of the possible recursive splits of a CU. Particularly, at the Intra prediction level, HEVC supports 33 angular prediction modes in addition to DC and planar modes, which is much more than the maximum of 8 angular modes proposed by H.264/AVC. Furthermore, the new standard allows deriving the “most probable mode” from neighbor blocks. In the case of the Chroma component, the same mode as the Luma can be used. Moreover, HEVC supports additional reference sample smoothing as well as a boundary smoothing.

These sophisticated prediction features offer a better coding efficiency, but at the cost of significant complexity at the encoder side. This complexity gives a special importance to developing fact mode decisions, especially for some applications and devices that do not support huge resources and that need to deal with real time encoding. In this context, we propose, in this paper, to investigate the potential of the gradient based Intra prediction.

A such gradient approach offers the possibility to estimate mathematically the gradient direction at each pixel position. This estimation is offering an interesting solution to take advantage of the large number of the angular Intra modes.

To generate the gradient information, the well-known Sobel operator is widely used in many video and image algorithms and applications. The reason behind this is the fact that this operator has one of the best edge detection performances over all the existing gradient operators.

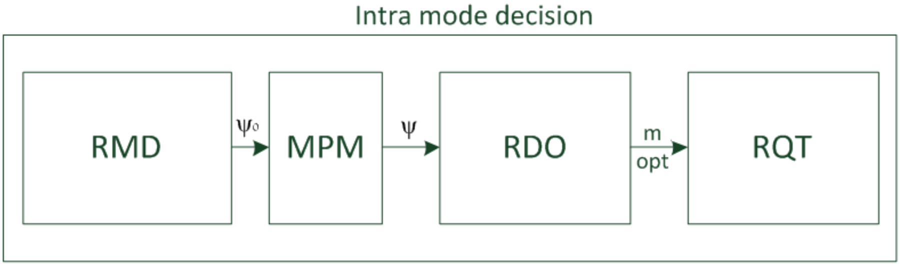

Fig. 2. Four stage Intra prediction

The related works that proposed a gradient stage to speed up the intra prediction use the 3-dimentonal Sobel operator like Pan et. al for H.264/AVC [3] and Jiang for HEVC [4]. In [5], we have investigated the potential of the Prewitt filter, presenting simpler coefficients than Sobel. In that work, we have generated a granular gradient map to reduce the preprocessing complexity coupled with a mode selection approach based on pixel neighbor extension.

In this work, we propose to investigate on HEVC Intra modes detection potential of the Roberts operator, a 2-dimensional operator as well as the complexity reduction that it can offer for hardware implementation. To be able to use such less accurate operator, we propose two approaches that allow enhancing the gradient decision on the Intra mode decision without introducing a significant complexity.

The remaining of this paper is structured as follows. The Section 2 presents an overview of the HEVC Intra prediction algorithm. Thereafter, Section 3 exposes the gradient based Intra prediction algorithm. We present then, in the Section 4 the proposed optimization approaches dealing with a pre-selection of Intra mode as well as an optimized mode selection at PU level, generated from the gradient information.

Section 5, then, exposes the approach of exploiting the gradient information to speed up the Intra prediction. Thereafter, the section6 presents an analysis of the complexity reduction offered by the Roberts operator. Then, Section 7 presents the experimental evaluation of the proposed algorithm. And finally, we present the conclusion in Section 7.

2 HEVC Intra Prediction

To speed up the Intra prediction, the HEVC test model (HM) [6], adopted a simplified Intra prediction algorithm which goes through 4 stages process for each PU [7] as presented in Figure 2. In the first stage, referred to as the rough mode decision (RMD), the HM performs a Hadamard Transform for each PU possible size, for all the 35 possible Intra modes, to generate the Sum of Absolute Transform Difference (SATD).

The SATD will be used in the estimation of the R-D cost of that PU, as shown in the following equation:

$$J = SATD + \lambda . R \tag{1}$$

where λ is Lagrangian multiplier and R is the bit consumption estimation.

The n best intra modes are taken to form the candidate set. The number of the candidate modes n, is set to 3, 3, 3, and 8 respectively for PU sizes of 64x64, 32x32, 16x16 and 8x8. At a second stage, a check is performed for additional MPMs (most probable modes) that are derived from neighbors, and are added to the candidate set if they are not already included. Then, a rate-distortion optimized quantization (RDO) is performed, at a third stage, for the modes of the candidate set at only the maximum size of TU, to pick the best intra mode for the PU as well as the best PU split structure at rate-distortion wise. In the last stage, for each PU, the best intra mode found previously is used in order to find the optimal Residual Quadtree (RQT) structure.

3 Gradient Based Pre-processing

To calculate the gradient, a discrete differentiation operator is used. The most widely used operator is the Sobel. This operator has two 3x3 kernels S_x and S_y shown in equations (2) and (3), used to approximate the horizontal and vertical derivatives of a two dimensions matrix.At each pixel position on the original image represented here as a matrix A, we perform a convolution through the two kernels, as shown in (4) and (5), to generate two matrices G_x and G_y which represent respectively an approximation of the horizontal and vertical derivatives at each pixel.

$$S_{x(Sob)} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \tag{2}$$

$$S_{y(Sob)} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{3}$$

$$G_x = S_x \times A \tag{4}$$

$$G_y = S_y \times A \tag{5}$$

$$\Phi = artan\left(G_y / G_x\right) \tag{6}$$

Equation (6) is then used to generate the corresponding gradient direction. In each pixel position, this direction points to the most important variation of pixels intensity.

That means that in the case of a pixel located on an edge, the gradient direction goes across that edge. For our case, we have to take the perpendicular direction to the gradient as it represents the similarity of pixels intensity. We simplify the direction computation presented in equation 6 to simply computing G_y/G_x. We define a look-up table with HEVC intra directions and correspondent G_y/G_x values. We pick from the look-up table the supported Intra direction that is the nearest to the obtained Φ value. And we affect the corresponding Intra mode to the current pixel location.

For complexity reduction, the gradient magnitudes can be roughly approximated by:

$$M = |G_x| + |G_y| \tag{7}$$

At the end of this pre-processing step, we will have a matrix of modes where each mode corresponds to a pixel location in the original picture. We mention here that the generated mode matrix contains only angular modes. DC and planar modes are not represented and as these two modes have great probability to be the best modes at the end of the Rate-Distortion evaluation, we include them automatically in the candidate set.

To reduce the complexity of the convolution calculation, a process that is generated at each pixel position, we propose to use here the Roberts operator.

$$S_{x(Rob)} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{8}$$

$$S_{y(Rob)} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{9}$$

As shown in the equations 8 and 9, the Roberts operator is based on 2x2 kernels that allow generating diagonal differences. So for the implementation, the gradient direction becomes:

$$\Phi_{Roberts} = artan(G_y/G_x) + \pi/4 \tag{10}$$

The smaller kernel size of the operator is offering an option to more simplify the gradient algorithm and so motivates our investigation in this work. But to enhance the gradient based solution and especially when using Roberts, less accurate operator than the wildly used Sobel, we propose in the next section some approaches that would improve the gradient performance.

4 Optimization of Intra Mode Detection

4.1 Optimal Mode Selection

Jiang has considered the accumulated gradient magnitudes $\mathrm{M_m}$ for each mode in the current PU, as a criterion to choose the best modes for the candidate set:

$$\mathrm{M_m} = \textstyle\sum_{\mathrm{i \in PU}} \mathrm{M_{m,i}} \tag{11}$$

where $\mathrm{M_{m,i}}$ is the gradient magnitude of a point i that have the mode direction m.

We propose, in this work, to take into account one more factor in the selection, which is $\mathrm{N_m}$, the number of appearance of a mode in the current PU. In fact, we can have, in some cases, a mode that appears in many points in the PU but with small magnitudes representing a spread variation of pixel intensity but with very small values.

And we can have, in other cases, a mode that exists in few points but with high gradient magnitudes reflecting a limited but high variation of pixel intensity. So as both the most appearing modes and the modes with high gradient values would approach the optimal Intra mode, we investigate the following cost function for each angular mode m:

$$\mathrm{Cost}_{\mathrm{m}} = \mathrm{N}_{\mathrm{m}} + \sum_{\mathrm{i}\in\mathrm{PU}} \mathrm{M}_{\mathrm{m,i}} \tag{12}$$

4.2 Mode Pre-selection

As the generated modes are just approximations, we propose in this section to extend, at each pixel position, the detection of a mode m, to consider two more angular modes if they exit. These two modes are the neighbor modes m+1 and m-1 of the detected mode m.

To farther investigate this extension, we express the cost function $\mathrm{Cost}_{\mathrm{m}}$ considering $\mathrm{Cost}_{\mathrm{m,i}}$, the cost of a gradient point i in the current PU which correspond to the Intra mode m :

$$Cost_m = \sum_{i\in PU} Cost_{m,i} \tag{13}$$

As expressed in the equation (15), for each detected mode m, we increase the $\mathrm{Cost}_{\mathrm{m,i}}$ by a bonus value b_{m} used to favor the detected mode against its neighbors. Similarly, the $\mathrm{Cost}_{\mathrm{m-1}}$ and $\mathrm{Cost}_{\mathrm{m+1}}$ of the neighbor modes $\mathrm{m}+1$ and $\mathrm{m}-1$ are increased by a neighboring bonus value b_{n}, used to favor the two neighbor modes against the other modes. For investigation on the best bonus values, we consider for different values of b_{m} and b_{n} the percentage of matching the best theoretical best mode in the candidate set for each PU.

$$\begin{cases} Cost_{m-1} = (1 + M_{m,i}) \times b_n \\ Cost_{m,i} = (1 + M_{m,i}) \times b_m \\ Cost_{m+1} = (1 + M_{m,i}) \times b_n \end{cases} \tag{14}$$

where b_{m} and b_{n} are the used bonus values so that $\mathrm{b}_{\mathrm{m}} > b_{\mathrm{n}}$.

As we noticed favorable results for the bonus values (b_{m}; b_{n}) equal to (3;2), so in the remaining of this paper, we continue working with these bonus values.

5 Fast Mode Decision

As mentioned before, all the 35 modes will be tested, in the RMD stage through a Hadamard Transform encoding in order to choose the best modes for the current PU. The idea here relies on a selection of the most probable modes, in order to limit the number of Intra modes to be tested and so speed up the Intra prediction process. In fact, the gradient generated histogram, for each PU, includes the costs $Cost_m$ of each of the Intra modes.

These costs will reflect a kind of probabilities for the modes to match the optimal mode for the current PU. Therefore, instead of going through all the modes, only a limited list of modes will be investigated. We refer to this list as the gradient candidate set, ψ_i^G where $0 \leq \mathrm{i} \leq N_G$, N_G being the number of modes in the candidate set. The gradient modes are ordered from most probable to least probable.

The gradient generated modes are more precise for higher sizes of PU as it has more points to approximate the most representative gradient of the PU. Thus, the number of modes N_G has to be set accordingly. We set this number to 15, 14, 8, 6 and 5 for respectively PU sizes of 4x4, 8x8, 16x16, 32x32 and 64x64, as we noticed that under theses settings, we have good tradeoff between time saving and encoding performance.

The best modes obtained through the RMD will form the RMD candidate set referred to as ψ_i^R, where $0 \leq \mathrm{i} \leq N_R$, N_R being the number of modes. We keep the number of modes N_R as it set in HM12.0, i.e. 8, 8, 3, 3 and 3 for respectively PU sizes of 4x4, 8x8, 16x16, 32x32 and 64x64.

In order to speed up the RDO process, the heaviest stage in the Intra prediction, we propose to reduce even more the number N_R for PU sizes of 8x8 and 4x4, based on the gradient stage performance of detecting the theoretical optimal mode. So, we reduce the number N_R according to different confidence scenarios.

These scenarios are set by comparing the candidate set ψ_G, result of gradient stage, to the candidate set ψ_R result of the RMD stage. The idea relies on the hypothesis that the more the results are similar, the more the gradient stage is approaching the theoretical optimal mode. So if the best RMD mode and best gradient mode are neighbors, we reduce the tested modes to only the best RMD one:

$$\text{if } |\psi_0^G - \psi_0^R| \leq 1 \quad \text{then } N_R = 1 \tag{15}$$

6 Complexity Analysis

We analyze in this section, the complexity of the implementation for both operators. The implementation difference concerns the convolutions process expressed for both operator in the equations below:

$$G_{x(i,j)}^{Sobel} = A_{(i+1,j-1)} - A_{(i-1,j-1)} + (A_{(i+1,j)} - A_{(i-1,j)}) \ll 2 + A_{(i+1,j-1)} - A_{(i-1,j-1)} \tag{16}$$

$$G_{y(i,j)}^{Sobel} = A_{(i-1,j-1)} - A_{(i-1,j+1)} + (A_{(i,j-1)} - A_{(i,j+1)}) \ll 2 + A_{(i+1,j-1)} - A_{(i+1,j-1)} \tag{17}$$

$$G_{x(i,j)}^{Roberts} = A_{(i-1,j-1)} - A_{(i,j)} \tag{18}$$

$$G_{y(i,j)}^{Roberts} = A_{(i-1,j)} - A_{(i,j-1)} \tag{19}$$

The multiplication for Sobel can be replaced by simple binary shifts. Despite this simplification, from software complexity wise, the convolution with Sobel operator needs 14 operations for each pixel position. However, the Robert operator needs only 2 operations, which is much less than the Sobel one.

In addition to the reduced software complexity offered by Roberts, this operator is a quite interesting solution hardware-wise.

In fact, in addition to the fact that Roberts based solution presents much less instructions, it offers some key points which make it even more interesting and by far a more friendly hardware solution.

- Data loading: For one gradient point, the Roberts based solution needs two pixel lines loading while the Sobel operator solution needs 3 lines loading.
- Line-based data: For the convolution, the Roberts based solution needs only a rotation instruction to be able to apply multi data-subtraction. However, to benefit from the multi data instructions, the Sobel based solution needs to convert the line-based loaded data to column-based one, which is a heavy process for the implementation.
- Coefficients: the 2 and -2 coefficients, in the Sobel kernels, make the convolution implementation need to apply additional masks to isolate the pixels to be multiplied by these coefficients and also need to apply then extra addition/subtraction instructions.

7 Experimental Results

For performance evaluation, the proposed algorithm was integrated in HM 12.0, and simulations were performed conforming to common test condition specified in [8].

To compare the time effect of the algorithm, we consider the time gain:

$$\Delta T = (T_{HM12} - T_{Prop})/T_{HM12} \tag{20}$$

where $\mathrm{T_{HM12}}$ is the encoding time of HM12.0 and $\mathrm{T_{Prop}}$ is that of the proposed solution integrated on HM12.0. As the implemented feature concerns mainly the intra coding, we present the results for an All Intra (AI) coding for 8 bit depth coding.

As cited previously, we have set the number of modes in the candidate set to be tested in the RMD to 15, 14, 8, 6 and 5 for respectively the PU sizes of 4x4, 8x8, 16x16, 32x32 and 64x64 and for the RDO, we kept these numbers as defined in the HM12.0 (8, 8, 3 , 3 and 3 accordingly).

In the table 1, we present the Bjontegaard Delta rate (BD-rate) [9] measurement and time saving performance of the proposed gradient solution over that of HM12.0. We can see from the table, that the proposed Intra perdition algorithm provides a time reduction for all the sequences with an average value of 31.9% with an average increase in BD-rate of only 1.1%.

For better evaluation of the proposed features, we present in the table 2 a comparison with the Jiang work [4]. Also, in order to evaluate the proposed features, we consider 3 configurations of the proposed algorithm with different combinations of the proposed features. We obtained almost the same BD-rate performances as in [4] but with some difference in time reduction.This difference is related to the fact that Jiang has used different mode numbers in candidate sets for both rough mode decision as well as the rate distortion optimization stage. In this work, we have chosen the RMD mode numbers according to the HM mode numbers of RDO candidate set [7].

Table 1. Comparison of best mode matching mode

Class / Sequence		Y	U	V	ΔT
Class A	Traffic	1.0	0.0	-0.5	30.7
	PeopleOnStreet	1.1	-1.0	-0.1	29.6
	Nebuta	0.2	0.3	0.2	29.7
	SteamLocomotive	0.1	-0.5	0.0	31.2
Class B	Kimono	0.3	0.1	0.0	34.7
	ParkScene	0.7	-0.6	-0.6	33.8
	Cactus	1.3	-0.1	0.2	31.5
	BasketballDrive	1.7	1.1	0.5	34.1
	BQTerrace	0.9	-0.5	-0.3	34.2
Class C	BasketballDrill	1.0	0.4	-0.1	29.8
	BQMall	1.3	-0.5	0.0	29.3
	PartyScene	1.3	-0.9	-0.8	30.6
	RaceHorses	0.9	-0.4	-0.9	29.9
Class D	BasketballPass	1.2	-0.2	-0.2	37.3
	BQSquare	1.6	0.1	0.0	33.4
	BlowingBubbles	1.3	-0.6	-0.7	29.3
	RaceHorses	1.4	-0.4	0.2	29.2
Class E	FourPeople	1.1	0.0	-0.4	33.7
	Johnny	1.3	0.0	-0.1	33.5
	KristenAndSara	1.4	0.1	-0.8	33.4
Ave.		1.1	-0.2	-0.2	31.9

The first proposed configuration presents the gradient solution using Roberts operator instead of Sobel. The second configuration deals with a Robert based gradient solution combined with the mode decision optimizations. And the third configuration includes, in addition to the former cited features, the early RDO option. We can see from the table 2 that the basic Sobel gradient algorithm achieves an average time reduction of 11.7% with an increase of 0.6%in BD-rate.

The first proposed configuration offers almost the same time reduction but with 0.8% as an increase in the BD-rate. Such configuration shows the impact in coding efficiency of using a less accurate operator. However the second configuration, while offering also almost the same time reduction as the two former configurations, allows just 0.4% increase in BD-rate.

Table 2. Performance comparison

Class	[4] (Sob.)		Prop. (Rob.)		Prop. (Rob., opt.)		Prop. (Rob., opt., Fast Intra)	
	Y	ΔT	Y	ΔT	Y	ΔT	Y	ΔT
A	0.4	9.6	0.5	9.1	0.2	9.1	0.6	30.3
B	0.5	12.9	0.9	12.2	0.6	12.1	1.0	33.7
C	0.5	9.6	0.8	10.0	0.2	9.8	1.1	29.9
D	0.8	12.4	0.9	12.4	0.4	12.7	1.4	32.3
E	0.7	14.0	1.1	13.7	0.5	13.3	1.3	33.5
Ave.	0.6	11.7	0.8	11.5	0.4	11.4	1.1	31.9

So the optimizations on the mode decision allows the Robert solution to make up the precisions difference and even offers 0.2% in BD-rate, better than the Sobel basic solution. The third configuration, including early RDO option, allows to reach 31.9% as an average time reduction with 1.1% in BD-rate. This result confirms the interesting option of exploiting the gradient information in order to speed up the Intra prediction algorithm with a small loss in BD-rate. But what makes this solution quite interesting is that it achieves better performances while offering important potential for a hardware complexity reduction.

We precise here that the time execution profiling computed as shown in equation 22 aims to estimate the complexity reduction achieved by the gradient based algorithms compared to the Hadamard transform based prediction and not for comparing the two operators. This is due to the fact that the pre-processing stage is about 2% of the whole Intra encoding. We precise that further investigation can be done on the number of modes in the candidate sets to optimize it for better trade-off of time gain/BD-rate loss for the proposed solution.

8 Conclusions

In this paper, we have presented a pixel-based gradient Intra prediction for HEVC. The proposed algorithm uses the Roberts operator as a discrete differentiation operator in order to approximate the gradient of the concerned block in the original picture.

The algorithm generates a preferred direction for each pixel in the block, from which we select a candidate set of modes to be tested in a Rate-Distortion optimization level. The mode election can be optimized through neighbor mode extension and adapted cost function to take into account both the most appearing modes and those with higher gradient magnitudes. A comparison with the Hadamard transform based algorithm used in HM12, shows that the proposed algorithm achieves a time saving of 31.9% with an average increase in BD-rate of just 1.1%.

References

1. Sullivan, G.-J., Ohm, J.-R., Han, W.J., Wiegand, T.: Overview of the high efficiency video coding (HEVC) standard. IEEE Transactions on Circuits and Systems for Video Technology **22**(12), 1649–1668 (2012)
2. Lainema, J., Bossen, F., Han, W.J., Min, J., Ugur, K.: Fast mode decision algorithm for intra prediction in H.264/AVC video coding. IEEE Transactions on Circuits and Systems for Video Technology **15**(7), 813–822 (2005)
3. Pan, F., Lin, X., Rahardja, S., Lim, K., Li, Z., Wu, D., Wu, S.: Fast mode decision algorithm for intra prediction in H.264/AVC video coding. IEEE Transactions on Circuits and Systems for Video Technology **15**(7), 813–822 (2005)
4. Jiang, W., Ma, H., Chen, Y.: Gradient based fast mode decision algorithm for intra prediction in HEVC. In: International Conference on Consumer Electronics, Communications and Networks (2012)
5. BenHajyoussef, A., Ezzedine, T., Bouallegue, A.: Fast Gradient Based Intra mode decision for High Efficiency Video Coding. International Journal of Emerging Trends & Technology in Computer Science **3**(3), 223–228 (2014)
6. HEVC reference model. http://hevc.hhi.fraunhofer.de/svn/svnHEVCSoftware/
7. Zhao, L. Zhang, L., Zhao, X.: Further encoder improvement of intra mode decision. In: JCT-VC Meeting, Doc., vol. D283 (2011)
8. Bossen, F.: Common HM test conditions and software reference configurations. In: JCT-VC Meeting, Doc., vol. L1100 (2013)
9. Bjontegaard, G.: Calculation of average PSNR differences between R-D curves. In: ITU-T VCEG 13th Meeting, Doc. VCEG-M33 (2001)