SPEEDING UP HEVC INTRA CODING BASED ON TREE DEPTH INTER-LEVELS CORRELATION STRUCTURE

Thaísa L.da Silva[1,2], *Luciano V. Agostini*[3], *Luis A. da Silva Cruz*[1,2]

[1]Instituto de Telecomunicações – Coimbra, Portugal
[2]Department of Electrical and Computer Engineering – University of Coimbra, Portugal
[3]Group of Architectures and Integrated Circuits – Federal University of Pelotas, Brazil

ABSTRACT

The encoding efficiency of the High Efficiency Video Coding (HEVC) standard is considerably higher than that of its predecessors. The intra coding method adopted by this new standard includes a larger number of prediction directions than those of its preceding standards namely H.264/AVC. This improvement contributed to enhance the encoding efficiency, but carried with it a much larger computational complexity. This paper presents an algorithm to speed-up the intra mode decision procedure through the exploitation of inter-level correlation observed between the predominant edge orientation of the current PU and those of the PUs already coded at lower levels of the encoding quadtree. Experimental results show that the proposed algorithm provides a decrease of up to 39.6% in the HEVC intra coding processing time, with a small increase in bit-rate (avg. 1.2%) and a negligible reduction in PSNR values.

Index Terms— Intra coding, High Efficiency Video Coding (HEVC), hierarchical structure, inter-levels correlation, computational complexity

1. INTRODUCTION

Even though current H.264/AVC [1] encoders are already able to reduce bit-rates by 50% when compared to predecessors [2], recent high resolution video applications and multimedia services require even higher efficiency in video compression at much lower bit-rates [3]. To satisfy these requirements the emerging High Efficiency Video Coding (HEVC) standard [4] has been developed by the Joint Collaborative Team on Video Coding (JCT-VC) with the mandate of doubling the compression efficiency achieved by current state-of-the-art H.264/AVC standard with a negligible increase in computational complexity.

Among other improvements, when compared to H.264/AVC, HEVC intra coding defines up to 33 prediction directions from which the best one (in rate-distortion sense) is chosen by exhaustive search, a highly time-consuming procedure. This paper presents an algorithm to speed-up the intra coding mode decision, which exploits the inter-level correlation introduced by the HEVC hierarchical structure. Such algorithm uses of the predominant orientation of the texture of the PU to be encoded and analyzes the high degrees of similarity between the dominant edge orientation of the texture of the current PU and those of the PUs at previous tree depth levels.

The remainder of this paper is organized as follows. Section 2 presents an overview of the hierarchical structure used in HEVC as well as its intra coding tool. Section 3 presents the details of the proposed intra mode decision algorithm. Experimental results and comparisons with related works make up Section 4 and Section 5, respectively. Finally, the conclusions are presented in Section 6.

2. HEVC HIERARCHICAL STRUCTURE AND INTRA CODING OVERVIEW

In the HEVC standard, each video frame is divided into a number of square blocks of equal size called coding tree blocks (CTBs). Each CTB can be partitioned into one or more coding units (CU) using a hierarchical quadtree structure, where the CTBs are the roots of each coding tree. The encoder determines the best segmentation of the quadtree structure based on Rate-Distortion Optimization (RDO) techniques [5], which test and evaluates all possible partitions of the coding tree.

The CU dimensions can vary from 8x8 up to the CTB dimensions and each leaf CU is also further partitioned into prediction units (PUs). In the case of intra coded CU the partition can consist of one 2Nx2N PU or four NxN PUs, where N is half of the CU size.

The HEVC standard also defines another type of unit, the transform unit (TU), which is used to represent the blocks of prediction residues that are transformed and quantized. When the transformation process of the prediction residue is started, each CU is assumed to be the root of another quadtree structure called residual quadtree (RQT) which leaves are the TUs.

An example of the hierarchical quadtree structure used in HEVC is presented in Figure 1, where each 64x64 CTB is composed of one 64x64 CU, which in turn can be divided into smaller CUs.

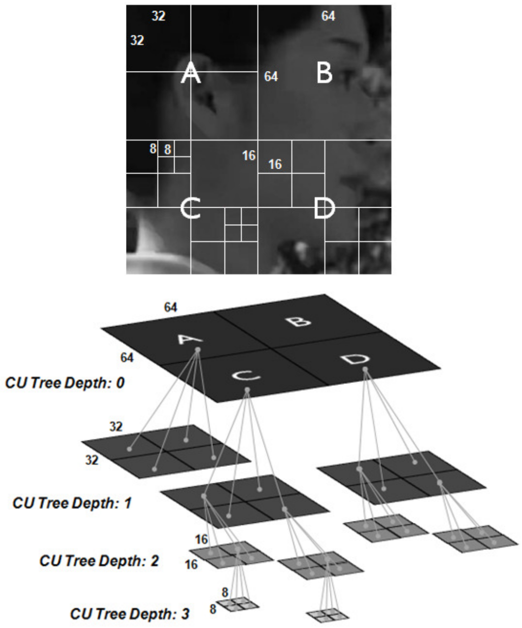

Fig. 1. HEVC CUs quadtree hierarchical structure

It is clear from the Figure 1 that the number of different trees to be tested in the RDO process increases quickly (exponentially with base 4) with the maximum depth of the tree.

The intra prediction procedure used in the HEVC provides a total of 33 intra prediction directions, as shown in Figure 2 [4], as well as two additional prediction modes: DC and planar [4]. To alleviate the computational burden of intra coding, the HEVC test models since HM 1.0 [6] uses a rough mode decision (RMD) process to reduce the number of intra modes to be evaluated during the final RDO stage.

As shown in Figure 3 this process starts by selecting a subset of candidate prediction modes composed of the modes that resulted in the smallest RD cost computed using the sum of absolute Hadamard transformed prediction residue (SATPR) and the corresponding bit-rate [7].

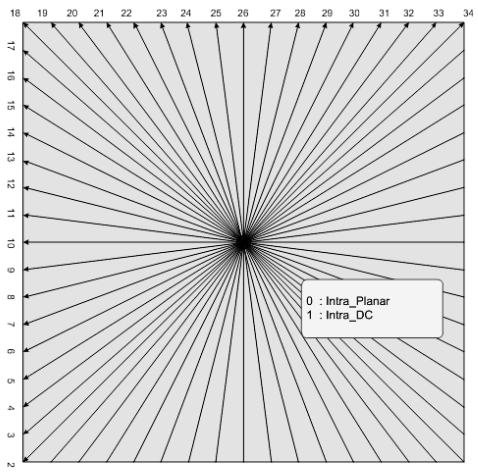

Fig. 2. Intra prediction mode directions in HEVC

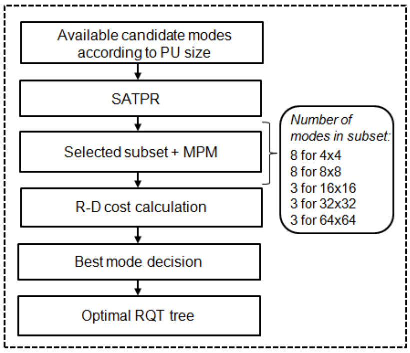

Fig. 3. Intra prediction using rough mode decision process

The size of the subset is chosen according to the PU size so that for 4x4 and 8x8 PUs it includes the eight best performing modes while for 16x16, 32x32 and 64x64 PUs it contains only the three best performing modes.

In the next stage, the R-D cost of each prediction mode belonging to the subset is computed and the best performing mode is selected to encode the PU. The definition of optimal RQT to be used for the encoding the residue obtained using the selected intra prediction mode concludes the process.

3. PROPOSED ALGORITHM

The proposed algorithm analyzes the correlation between the edge information of the current PU and the edge information of the PUs already coded at lower tree depth levels. This analysis is performed for all PU sizes (64x64, 32x32, 16x16, 8x8 and 4x4 PUs).

Initially, the proposed algorithm analyses the PU texture to determine its dominant edge orientation, which is then used to choose the subset of modes that will be used in the intra coding. Five edge orientation indicators are evaluated: horizontal, vertical, 45° diagonal, 135° diagonal and non-directional [8].

The edge orientation of each 4x4 PU is calculated based on the luminance values of its pixels by first dividing it into four 2x2 blocks: Block 0, Block 1, Block 2 and Block 3, as illustrated in Figure 4, and then computing each of these 2x2 blocks average values, c0, c1, c2 and c3, respectively.

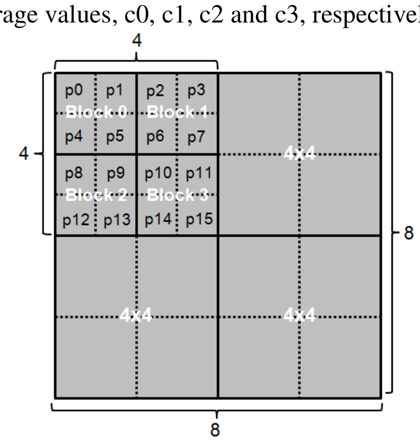

Fig. 4. 8x8 PU edge computation

These average values are then used in equations (1) to (5) [8] to compute the five orientation indicator strengths, eV, eH, e45º, e135º and eND (non-directional).

$$eV = |c_0 - c_1 + c_2 - c_3| \quad (1) \qquad eH = |c_0 + c_1 - c_2 - c_3| \quad (2)$$

$$e45^\circ = \left|\sqrt{2} \times (c_0 - c_3)\right| \quad (3) \qquad e135^\circ = \left|\sqrt{2} \times (c_1 - c_2)\right| \quad (4)$$

$$eND = |2 \times (c_0 - c_1 - c_2 + c_3)| \quad (5)$$

The edge orientation is determined from these edge strengths, as that for which the orientation indicator has the maximum value as per equation (6).

$$S = \underset{Or \in \{V, H, 45^\circ, 135^\circ, ND\}}{\arg\max} \{eOr\} \qquad (6)$$

For PUs larger than 4x4, such as the 8x8 PU in Figure 4, an average of the values for each edge direction is computed considering the edge orientation of all the 4x4 PUs that compose the larger PU, and the direction with the highest average is defined as the direction of the dominant edge of the bigger PU.

Five subsets of angular modes are defined, each one composed of nine angular modes taken from the full set of 33 angular modes defined in HEVC. Each subset is associated with one of the five edge orientations defined before (horizontal, vertical, 45°, 135° and non-directional), as presented in Figure 5.

The intra prediction process will then use only one of the five subsets selected according to the edge orientation computed from the pixels of the PU to be (intra) coded and only the modes belonging to that subset will be tried during the RDO search for the best prediction mode.

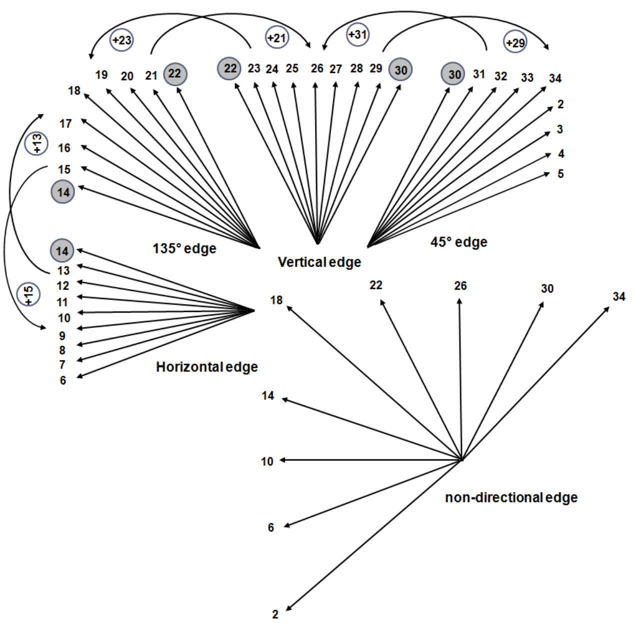

Fig. 5. Subsets of modes for each edge direction

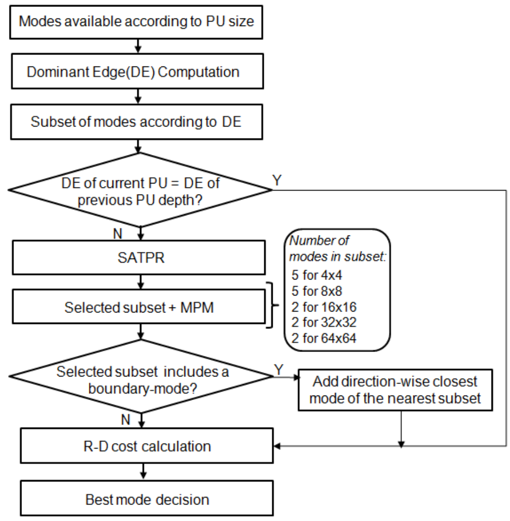

Fig. 6. Flowchart of the proposed algorithm

These subsets may be complemented with an additional mode, depending on the modes selected after to the SATPR calculation. After these modes have been chosen a check is made on whether any of the candidate modes selected corresponds to a boundary-mode, which is a mode that is at the boundary between two directionally neighboring subsets of modes. In Figure 5, the boundary modes 14, 22 and 30 are highlighted. If that condition holds true, the prediction mode directionally closest and belonging the nearest subset is added to the candidate modes subset. For example, if the 135° edge direction is computed as the PU dominant edge orientation and mode 22 (boundary-mode) was selected as a candidate mode (in step "Selected subset + MPM" Figure 6), mode 23 (vertical edge) is added to the subset of candidate modes to be tested in RDO process.

Since these subsets of modes selected according to the dominant orientation of the PU texture are much smaller than the original subset obtained after the SATPR procedure (even after addition of a boundary mode), their use in the prediction mode selection process is faster than the original method, as the results will reveal.

Subsequent to the PU edge evaluation, the test video sequences specified in [9] were analyzed, with quantization parameters 22, 27, 32 and 37, looking for correlations across adjacent tree level PUs. The experimental results are presented in Table 1 and show that in most cases the predominant orientation of the texture of the current PU is the same that of the PU at adjacent tree level PU.

Table 1. Inter-level correlation between the PUs orientation

QP	Same Orientation (%)	Different Orientation (%)
22	93.55	6.45
27	90.60	9.40
32	88.66	11.34
37	85.55	14.45
Average	89.59	10.41

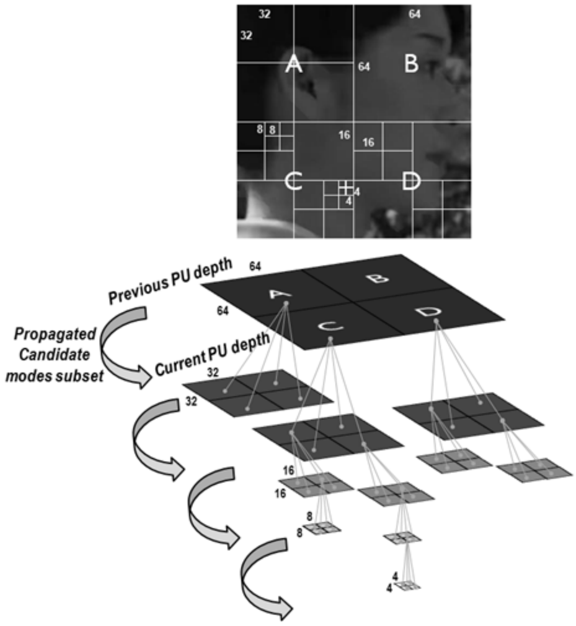

Fig. 7. HEVC PUs quadtree hierarchical structure

From the experimental results it was possible to verify that, in most cases, the dominant edge of the current PU has the same orientation than that of the PU at the immediately lower depth. This observation suggested the introduction of a new test to verify if the dominant edge of current PU is the same of the previous PU depth (fourth step in Figure 6).

If this is true, the candidate modes subset of the previous PU depth (as illustrated in Figure 7) is assumed to be a good choice for the current PU, the SATPR step is skipped and the R-D cost calculation is performed immediately, as shown in Figure 6 which presents the entire prediction mode decision algorithm.

4. EXPERIMENTAL RESULTS

To evaluate the computational complexity reduction of the proposed algorithm and its effect on encoding performance some tests were made using a modified version of the HEVC HM10 test model which included the method described so far. Since this work is focused on high resolution video applications, a group of experiments were carried out using six high-resolution test video sequences as specified in [9]. The test sequences were encoded with four quantization parameter (QP) values: 22, 27, 32 and 37, using the all intra-high efficiency configuration (AI-HE) [4].

Table 2 provides the detailed results for ΔPSNR, ΔBitrate, ΔTime and Bjontegaard deltas [10] [11] comparing the performance and complexity of the proposed algorithm with results obtained using the HM10 reference software. The proposed algorithm achieves an encoding time reduction of up to 39.6% in relation to HM10 (37% on average) with slight degradation in bit-rate (1.2% on average) and PSNR (0.07dB on average). Similar results were observed for lower resolutions video sequences tested.

Table 2. Comparison of proposed algorithm with HM10

Sequence	QP	ΔBitrate [%]	ΔTime [%]	ΔPSNR [dB]	BD-rateY [%]	BD-PSNRY [dB]
Nebuta Festival (2560x1600)	22	0.67	-36.07	-0.025	1.5	-0.11
	27	1.10	-36.39	-0.021		
	32	1.79	-37.44	-0.001		
	37	2.01	-36.85	-0.042		
Steam Locomotive (2560x1600)	22	0.41	-34.43	-0.024	1.7	-0.05
	27	1.09	-36.02	-0.033		
	32	1.52	-37.75	-0.060		
	37	1.75	-37.33	-0.068		
Kimono (1920x1080)	22	0.75	-34.57	-0.044	2.2	-0.07
	27	1.08	-36.29	-0.051		
	32	1.63	-38.29	-0.083		
	37	2.24	-36.92	-0.116		
ParkScene (1920x1080)	22	0.29	-34.15	-0.129	1.9	-0.08
	27	0.54	-38.68	-0.137		
	32	0.96	-38.37	-0.138		
	37	1.72	-35.03	-0.140		
SlideEditing (1280x720)	22	0.67	-36.54	-0.074	1.3	-0.18
	27	0.83	-39.60	-0.083		
	32	1.10	-38.83	-0.087		
	37	1.37	-37.99	-0.091		
ChinaSpeed (1024x768)	22	1.03	-36.07	-0.061	1.8	-0.16
	27	1.35	-39.43	-0.078		
	32	1.55	-38.96	-0.085		
	37	1.94	-38.32	-0.103		

5. COMPARISON WITH RELATED WORKS

In Zhao et al. [12] a rough mode decision (RMD) process is performed to reduce the number of intra mode candidates to be evaluated in the RDO procedure. Their work also exploits the strong correlation among the neighbors to identify a most probable mode, which is added to the candidate mode set. The work by Jiang et al. [13] takes a different approach; a gradient direction histogram is computed for each CU and based on this histogram a small subset of the candidate modes are selected as input for the RDO process. In Kim et al. [14] an early termination of intra prediction is performed based on the intra prediction mode used in the previous depth PU and on the block size of current depth TU. In addition, the number of candidates of the RMD is further reduced before the RDO procedure, and the intra mode of corresponding previous depth PU is always included in the candidates list for intra mode decision. Zhang et al. [15] propose an algorithm that analyzes the relation between PUs texture characteristics and from this analysis it adaptively reduces the number of candidate modes for the RDO process. In the method proposed by Zhang et al. [16] a 2:1 downsampled Hadamard transform is used for the rough mode decision. Furthermore, a gradual search to reduce the number of modes for Hadamard cost calculation and an early termination scheme is applied to speed up the RDOQ process. The method developed in Zhao et al. [12] has been partially adopted in HM version 2.0 (and followers). Since the algorithm proposed in this paper was tested on HM10 which also includes the procedure from Zhao et al. [12], the computation time savings reported in this paper add to those from Zhao et al. [12]. The method proposed in this work outperforms the computational complexity reduction reported by those authors [13], which reduces the intra prediction computation expenditure by 19.99% on average. Concerning encoding efficiency, the algorithm presented in Jiang et al. [13] incurred an average increase of 0.74% in bit-rate and average decrease of 0.04dB in PSNR values when compared to the HM4.0.

In terms of encoding time reduction the algorithm proposed in this work outperforms [14] which reports encoding complexity reductions up to 22.99% for the AI-HE case. When compared with Zhang et al. [15], the algorithm proposed in this work exceeds the reduction in complexity achieved by those authors, in which the encoding time can be reduced by up to 15% also for the AI-HE case. In performance terms the proposed algorithm by Zhang et al. [15] presented an increase in BD-Rate of 0.64% on average when compared to the HM4.0. When compared with Zhang et al. [16], the algorithm proposed in this work presents similar results in terms of complexity reduction achieved by those authors (38% on average). However, regarding encoding efficiency, the algorithm proposed in this work outperforms [16], which reports an increase in BD-Rate of 2.9%, while this work presented an increase of 1.7% on average.

6. CONCLUSIONS

This work presented an algorithm to speed-up the intra coding mode decision. Such algorithm exploits the hierarchical organization of the tree structure used in the HEVC standard, analyzing the inter-level correlation between the current PU predominant edge orientation and the predominant edge used in the PUs already predicted at previous tree depth levels.

Experimental results demonstrated that when compared with the HM10 intra coding algorithm, the encoding processing time was reduced by up to 39.6%, with a negligible drop in encoding efficiency, performing better than competing intra prediction speedup methods [12] [13] [14] [15]. An extension to this work is under way which is looking into methods to reduce the HEVC intra prediction complexity in the multiview scenario, by taking advantage of correlation of intra prediction modes across neighboring views.

7. REFERENCES

[1] G. Sullivan and T. Wiegand, “Overview of the H.264/AVC video coding standard,” IEEE Transaction on Circuits and systems for Video Technology, vol.13, no.7, pp.560-576, 2003.

[2] A. Puri, X. Chen, and A. Luthra, "Video Coding Using the H.264/MPEG-4 AVC Compression Standard," Elsevier Signal Processing: Image Communication, vol. 1, pp. 793-849, 2004.

[3] ITU-T Q6/16 Visual Coding and ISO/IEC JTC1/SC29/WG11 Coding on Moving Picture and Audio, “Call for Evidence on High-performance Video Coding (HVC),” N10553, 2009.

[4] T. Wiegand, W.-J. Han, B. Bross, J.-R. Ohm, and G.J. Sullivan, “High efficiency video coding (HEVC) text specification draft 8,” JCTVC-J1003, JCT-VC Meeting - Stockholm. July, 2012.

[5] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," IEEE Signal Processing Magazine, vol. 15, pp. 74-90, 1998.

[6] High Efficiency Video Coding Test Model. Available in: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-1.0/. 2012.

[7] 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, pp. 1792-1801, 2012.

[8] ISO/IEC/JTC1/SC29/WG11: “MPEG-7 XM Document: MPEG-7 Visual Part Experimentation Model Version 10.0.” MPEG Document N4063. Singapore. Mar. 2001.

[9] F. Bossen. “Common test conditions and software reference configurations”. JCTVC- K1100. JCT-VC Meeting. Shanghai. Oct. 2012.

[10] G. Bjontegaard. Calculation of Average PSNR Differences between RD curves. VCEG-M33. Apr. 2001.

[11] G. Bjontegaard. Improvements of the BD-PSNR model. VCEG-AI11. Jul. 2008.

[12] L. Zhao, L. Zhang, S. Ma, and D. Zhao, “Fast Mode Decision Algorithm for Intra Prediction in HEVC”, IEEE Visual Communications and Image Processing (VCIP), pp. 1- 4, 2011.

[13] W. Jiang, H. Ma, and Y. Chen. “Gradient Based Fast Mode Decision Algorithm for Intra Prediction in HEVC”, 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet), 2012, pp. 1836 – 1840.

[14] J. Kim. J. Yang. H. Lee, and B. Jeon. “Fast Intra Mode Decision of HEVC based on Hierarchical Structure”, 8th International Conference on Information. Communications and Signal Processing (ICICS), pp. 1-4, 2011.

[15] M. Zhang, C. Zhao, and J. Xu. “An Adaptive Fast Intra Mode Decision in HEVC”, International Conference on Image Processing (ICIP), pp. 221-224, 2012.

[16] Zhang, H., and Z. Ma. “Fast Intra Prediction for High Efficiency Video Coding”, in Pacific-Rim Conference on Multimedia (PCM), pp. 1-10, 2012.