CLUSTERING-BASED FAST INTRA PREDICTION MODE ALGORITHM FOR HEVC

Sami Jaballah [*] [†], *Kais Rouis* [*] [†], *Jamel Belhadj Tahar* [*]

[*] InnovCom Laboratory,
Higher School of Communications of Tunis, SupCom,
University of Carthage Tunis,Tunis, Tunisia

[†] Université de Tunis El Manar,
Ecole Nationale d'Ingénieurs de Tunis,
Tunis, Tunisie

ABSTRACT

The High Efficiency Video Coding (HEVC) is the next generation video coding standard.The HEVC provides equivalent perceptual quality with bit-budget saving greater than 50% compared to the H.264/AVC. In this paper, we propose a new fast intra prediction mode decision algorithm for the HEVC. We apply an early termination method based on the statistics of the resulted IPMs from both rough mode decision and most probable mode stages. The resulted IPMs are clustered into a K cluster by means of the K-medoid clustering algorithm, and each IPM cluster center represents all the IPMs within a cluster for the RDO process. The sugeested algorithm has been evaluated on high resolution test video sequences. Compared with the current HM16.0 and state-of-the-art scheme in all intra high efficiency configuration cases, the proposed algorithm outperforms the state-of-the-art scheme in terms of encoding time with similar coding efficiency

Index Terms— HEVC, intra prediction, mode decision, early termination

1. INTRODUCTION

The High Efficiency Video Coding (HEVC) is the latest video coding standard developed by the Joint Collaborative Team on Video Coding (JCT-VC) formed in April 2010 by ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG) [1]. The HEVC has been particularly designed to handle ultra-high definition resolution video applications and evidently, all existing applications of H.264/MPEG-4 AVC [1]. It also offers a coding efficiency improvement about 50% bit-rate reduction compared with the previous standard H.264/AVC.

In order to enhance its coding efficiency, and to to deal with higher video resolution such as $4K$ and $8K$ formats, the HEVC extends the concepts of the macro-block in H.264/AVC to a new structure called the Coding Tree Unit (CTU). The CTU is composed from Luma Coding Tree Blocks (CTBs), the corresponding Chroma CTBs and syntax elements. Each CTB can be split into one or many Coding Blocks (CBs), with a size that can be selected by the encoder as 16×16, 32×32, and 64×64. Otherwise, the Prediction Unit (PU) provide necessary information for prediction process. Based on the chosen prediction mode, each Coding Unit (CU) can be split into one, two or four Prediction Blocks (PBs). In general, five PB sizes are supported in the current HM16.0 for intra prediction, which are 4×4; 8×8; 16×16; 32×32; and 64×64. The HEVC encoder selects the best intra luma prediction mode using Unified Directional Intra (UDI) prediction that supports 35 modes, including two non-directional modes and 33 directional modes [1] [2]. The best direction is elected by choosing the minimum mode resulting from the Rate-Distortion Optimization (RDO) process. In order to reduce the computational time of the encoder and the number of candidate modes for the RDO, a Rough Mode Decision (RMD) applies to all candidate modes. The RMD determines the minimum absolute sum of Hadamard Transformed coefficients of residual signal (HSAD) and the mode bits. N modes are selected according to their RMD cost. The Most Probable Mode (MPM) proposed in [3], exploits the neighboring intra prediction information to reduce the number of directions taking part in the RDO process. However, the HEVC suffers from a high computation complexity and it is reasonable to consider the correlation between resulting the RMD and MPM intra prediction modes [3]. Our contribution comes with the following standoffs. Initially, we group the N best RMD mode candidates for the RDO process and the MPM modes if they are not included into a new intra prediction mode set. We introduce a new early termination criterion based on the statistic of the Hadamard cost of modes in the new set. We cluster the set of the intra modes into K clusters using the K-medoids clustering algorithm. We choose the K cluster centers to be the candidates for the RDO process. Finally, the refinement process is adopted by selecting all modes in the best cluster selected by the RDO.

The remainder of this paper is organized as follows. In section 2, we present the HEVC intra prediction mode decision process. The previous approaches introduced in the literature are mentioned in section 3. Our proposed intra prediction mode decision is fairly described in section 4. The experimental results and comparisons are provided in section 5. Finally, the contributions of this paper are summarized and the future work is outlined in section 6.

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

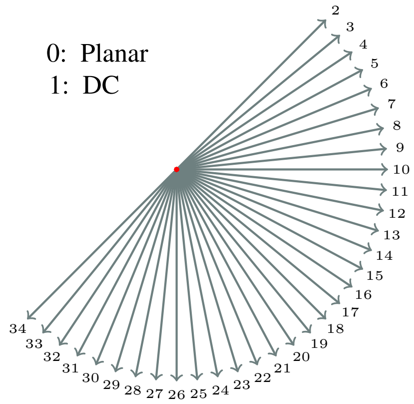

Fig. 1. Intra Prediction Modes in HEVC

2. THE HEVC INTRA PREDICTION

The intra prediction of a frame in a video sequence exploits the spatial redundancies within the same frame. The intra prediction major aims is to guarantee error resilience and temporal random access points in a video, more than providing high coding gain [4]. In the H.264/AVC, the intra prediction is performed using already encoded neighbouring macro-blocks (left, up and top-right). The HEVC adopts the same principle as in H.264/AVC, by using the boundary pixels to predict the target block. Furthermore, the UDI prediction mode decision in the HEVC provides up to 35 prediction modes, which include 33 angular prediction modes and two no-directional ones (DC and Planar), as shown in Fig.1. In the HEVC, an intra-predicted Coding Unit (CU) may be associated to two PU partition types which are : PART-2N2N and PART-NN. If the PART-NN prediction type is selected, the CU is split into 4 PUs. Otherwise, the CU maintains the same size.

After selecting the size of the target prediction block, the HEVC checks all available prediction modes for this size. The available number of intra modes depends on the size of the PU target: 4 modes for 64×64 PUs; 18 modes for the 4×4 PUs; and 35 modes for 8×8, 16×16 and 32×32. The prediction direction in intra prediction has the angles of +/- [0, 2, 5, 9, 13, 17, 21, 26, 32]/32. The set of intra angular prediction is designed to provide coverage for a region near horizontal, vertical and diagonal angles. As in H.264/MPEG-4 AVC, the HEVC supports two other prediction modes for the intra angular prediction.

A fast intra prediction mode decision is implemented in the current HEVC test Model HM16.0, which reduces the number of candidate modes for the RDO process. First, an RMD process is performed to obtain the best N candidates' list. These candidates are selected by computing the Sum of Absolute Transformed Differences (SATD) between the target and the predicted PU, where N candidates yielding the smallest cost are added to the list. The intra modes of the already encoded neighbouring blocks (MPM) are added to the

Table 1. Number of available intra prediction modes for each PU size

PU size	No. For RMD	No. For RDO
64×64	4	3+MPM
32×32	35	3+MPM
16×16	35	3+MPM
8×8	35	8+MPM
4×4	18	8+MPM

candidate list if they are not included. Finally, all N+MPM intra prediction candidate modes are evaluated by the RDO process, and the mode with the best R-D cost is selected to encode the block. The number of candidates for the RMD and RDO are presented in Table. 1.

3. RELATED WORKS

Recently, different algorithms have been proposed in order to reduce the computation of the mode decision for intra prediction, with a relative loss of coding efficiency. A new fast algorithm for HEVC intra coding is proposed in [5]. This algorithm proposes a depth range prediction method by exploiting the correlation between the neighbouring CTUs. To decide whether the current CU is split or not, the rate distortion costs and HSAD of recently encoded CUs are employed in this algorithm. As a final step, this algorithm selects only the IPMs with lower precision to be employed in the RMD, and the IPMs for the RDO are reduced based on the correlation between the neighbouring CUs. In [6], a gradient-based fast mode decision algorithm for the HEVC aims to reduce the computational complexity. By making use of the gradient information for each CU, a gradient-mode histogram is generated. Based on the histogram distribution, the number of candidate modes for the RMD and RDO processes is reduced. Based on edge direction, authors in [7] introduced a fast HEVC intra prediction mode decision approach. This latter calculates the predominant orientation of the edge of the current PU pixels. Using orientation information, it defines a reduced subset of intra modes to be evaluated from which the best mode is finally selected.

Otherwise, a group-based fast intra mode algorithm [8] reduces the computational complexity of the HEVC encoder. The method applies an early termination if the first RMD cost is greatly smaller than others. First, the rough modes are merged according to their angles and then, an edge detection process is applied to choose the best angle direction. The authors in [9] have recently put forward a fast intra prediction mode decision algorithm based on micro-level and macro-level schemes. The first scheme consists of a Hadamard cost-based progressive Rough Mode Search (pRMS) to check only potential modes. The pRMS selects few effective candidates to get the optimal rate distortion mode. To further reduce the complexity, an early rate distortion optimisation quantization skip method is adopted in this scheme. Afterwards, early CU-split termination is employed using the aggregated R-D cost of partial sub-CUs. Despite the fact that the above algorithms provide convincing results, the time saving or the computational overhead needs to be improved by further exploiting the static correlation between the IPMs Hadamard cost. Furthermore, an other possible solution would be effective for a light-weight algorithm by using an early termination criterion.

4. PROPOSED FAST INTRA PREDICTION ALGORITHM

A priori early mode selection termination is carried out to speed up the coding process in HEVC. Certain criteria must be satisfied for selecting the corresponding mode from the overall set as described in subsection 4.1. The rough decision mode used in HEVC chooses some prediction modes having close Hadamard costs, which yields to similar coding performance. One can bear in mind that conserving the corresponding coding efficiency while considering the computational complexity of a given encoding process, could be a chicken-egg dilemma for most mode decision approaches. We take into account this intrinsic relationship so that we adopt a clustering algorithm, explained in subsection 4.2.

4.1. Early Termination

For an adequate IPM selection, the HEVC evaluates the RD cost of the best RMD candidate modes and the MPM. Theses modes are ordered according to their Hadamard costs in an ascending order. Recent works on the HEVC intra prediction mode decision [8] [9], have assumed that the first two RMD modes have a high probability to be selected as the best mode. In this work, we apply the early termination scheme on both RMD and MPM modes. These modes are associated together in a one set, which are sorted out according to their Hadamard costs. The early termination criterion is activated if the following constraints are met :

- The two first modes in (RMD+MPM) set, namely ξ, are not angular neighbour modes or non-directional modes. $|M_1 - M_0| > T$, where T is a threshold to be determined.
- $Cost\,(M_1) - Cost\,(M_1) > \lambda$, where λ is the standard deviation of ξ
- $\frac{Cost(M_0)-Cost(M_{N-1})}{2} > \mu$, where N and μ are respectively the size and the average of ξ.

The threshold plays an important role in the activation of the early termination process. In this work, the values of the threshold T are taken between 2 and 4, otherwise the two modes are not considered angular adjacent modes.

4.2. K-medoids based intra prediction mode decision

The main idea consists in grouping generated all available modes (RMD+MPM) into statistically homogeneous groups, where each group is represented to the RDO by a single mode. In fact, the RDO candidate modes could be further reduced. Therefore, we based on a clustering technique using the K-medoids algorithm to group modes and select the best representative one, within each cluster. Moreover, all IPMs (RMD+MPM) are considered as the input data for the K-medoids algorithm. After performing the clustering steps, K IPMs modes are selected to replace the remaining ones and therefore, the RDO process selects the best mode. In the same group of this latter, the algorithm performs subsequently another RDO process as refinement mode step for the remaining modes. But this step sometimes cannot achieves a better coding gain in the case where the number of RDO candidate modes has not enough reduced. So we restrict the refinement mode step for every prediction bloc that has an available number of IPMs for the RDO lower than $2 * K$.

Clustering analysis is an important technique in the field of data analysis. The main goal of clustering is to maximize the homogeneity between objects within a cluster and the heterogeneity between different clusters. Many distance metrics are used to quantify the degree of dissimilarity among patterns. Each distance is more or less suitable to a given application. In the same vein, only unlabeled data and the number of partitions are provided. The clustering can be considered as an unsupervised classification technique.

K-Medoids algorithm selects the most centrally located object of a cluster as the cluster center [10]. Unlike the K-means algorithm, the K-medoids takes a representative object (medoids) as a center in each cluster instead of using the mean value of objects (non-medoids) within the cluster . The non-medoids objects are associated to the nearest cluster with respect to their distances from the medoid object. A swap operation is performed between actual medoid and each non-medoid object within the corresponding cluster in order to configure the medoids' distributions. Once new medoid objects are maintained, the hole data repartition get updated according to the new configuration. Details concerning the main functionalities of the algorithm are described below:

In the next section, we present simulation results and we compare the proposed algorithm using the HM16.0 as a reference software, with a state-of-art algorithm [7]. We will show its significant decrease of up to 52.1% in the coding computing time.

5. RESULTS AND DISCUSSION

We consider in our experimental tests the coding performance and computational complexity for the evaluation performance. The algorithms were executed using six well-known video sequences, with the quantization parameters 22,

Algorithm 1: K-medoids

1 Choose randomly K medoids from data.

2 Assign each object to the clusters with the closest medoid using the Minkowski distance.

$$d_{i,j}(X_i, M_j) = \left(\sum_{l=1}^{D} |x_{i,l} - m_{j,l}|^p \right)^{1/p} \quad (1)$$

$d_{i,j}$ denote the Minkowski distance between data vector X_i and medoid vector M_j and D is the vector dimension.

3 For each medoid $m = 1, ..., K$ and non-medoid data $x = 1, ..., n$
- Swap m and x and compute the total cost of the configuration.

4 Select the configuration with the lowest cost.

5 Repeat Steps 2 to 4, until the stopping criterion is met.

27, 32 and 37.

The test was performed under common test conditions (CTC) [11] and all-intra high-efficiency configuration cases. The platform is Intel Core i3-2350M CPU @2.30GHzx4 processor with 3.8 GiB RAM and Ubuntu 12.10. In Table 2, we list the comparative results of the test video sequences provided from our proposed algorithm and the anchor HM16.0. The number of clusters for the K-medoids algorithm is fixed to $K = 3$ and the early termination threshold $T = 3$. The comparison criteria used in this work are the Bjontegaard metrics; BD-PSNR and BD-Rate [12]. In addition to these metrics, we used the average percentage difference in coding time (ATS) to compare our method with the intra mode decision algorithm implemented in the HM16.0.

$$ATS = \frac{1}{4} * \left(\frac{Time_{proposed} - Time_{HM16.0}}{Time_{HM16.0}} * 100 \right) \quad (2)$$

Table 2 shows the coding efficiency and the time saving of the proposed algorithm over the current HM16.0 in cases when the early termination is adopted in the algorithm or not(ET-ON or ET-OFF). If we consider k-medoids algorithm

Table 2. Performance comparison of the test squences

	ET-OFF			ET-ON		
Sequences	ATS	BD-R	BD-P	ATS	BD-R	BD-P
Traffic	-17.96	0.9	-0.04	-32.96	0.8	-0.05
PeopleOnStreet	-35.6	0.9	-0.05	-52.1	0.8	-0.07
ParckScene	-16.74	1.0	-0.04	-29.54	0.9	-0.05
Kimono	-19.65	0.9	-0.03	-33.21	0.8	-0.04
PartyScene	-18.7	1.0	-0.08	-16.46	0.9	-0.1
RaceHorcesC	-27.42	0.9	-0.02	-49.3	0.9	-0.03
Average	**-22.67**	**0.95**	**-0.05**	**-35.59**	**0.85**	**-0.08**

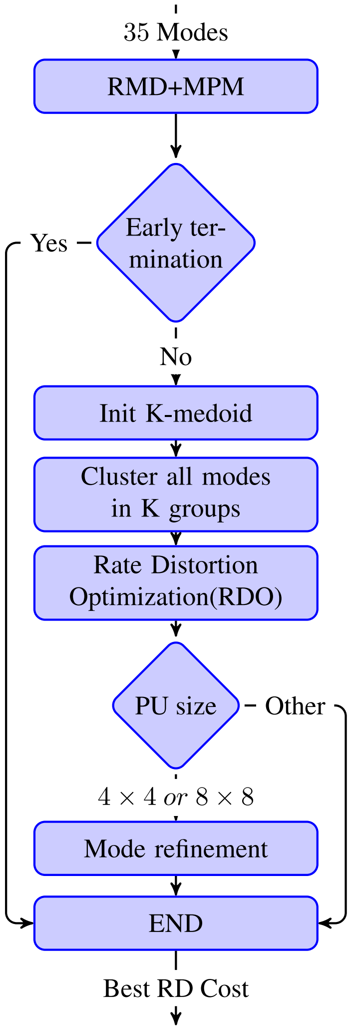

Fig. 2. Flowchart of the proposed algorithm

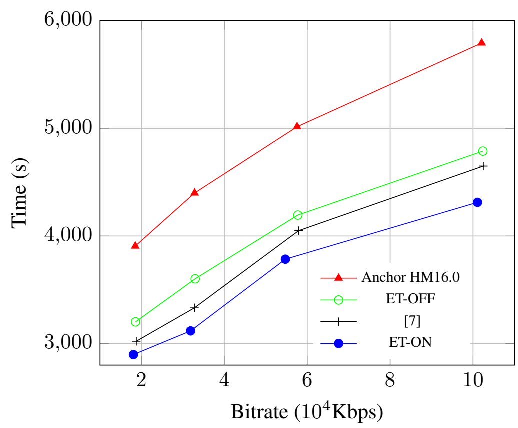

Fig. 3. Encoder time comparison of the HM16.0, proposed algorithm(ET-OFF and ET-ON) and [7] over "traffic 2560x1600".

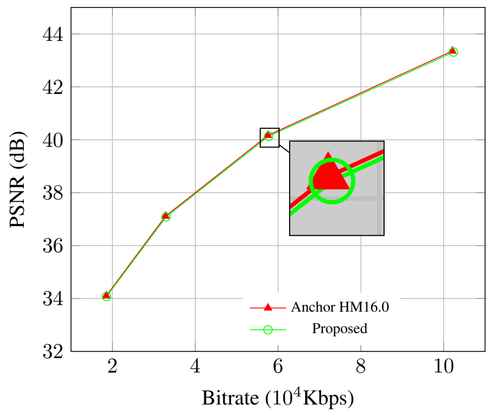

Fig. 4. RD curves of "Traffic 2560x1600".

with ET-OFF, an interesting outcome from these results is that the novel technique outperforms the HM16.0 with an average time saving about 22%. Otherwise, our method has RD performance loss about 0.95 for BD-Rate and 0.05 for BD-PSNR. For the $PeopleOnStreet$ and $RaceHorces$ video sequences, we obtain an average time saving about 35% and 27%, respectively. Using the early termination scheme (ET-ON), our algorithm shows a significant improvement in terms of time and bit-budget saving, with a slightly PSNR degradation. As observed for the previous case (ET-OFF), the highest ATS is obtained for the same two video sequences with 52.1 and 49.3, respectively.

Fig.3 shows the encoding time of our method, the ET-OFF, ET-ON, method in [7] and HM16.0 for the $Traffic$ sequence. We reached a significant time saving when the early termination criterion is adopted. Fig.4 illustrates the RD performance of our proposal over the HM16.0 for the $Traffic$ sequence. Results further indicate that the RD-curves of our algorithm and the HM16.0 are almost superposed, which means that the distortion is similar at the same bit-rate.

6. CONCLUSIONS

We have described a novel HEVC intra prediction mode decision aiming to reduce the HEVC encoder computation time. An early termination criterion is developed for this algorithm based on statistics of Hadamard cost of the available modes. The k-medoids clustering algorithm is therefore used to cluster into K partitions all available modes for RDO. Based on the empirical comparison results of our solution against [7] and the HM16.0 anchor over six test video sequences, a significant time encoding reduction and negligible performance loss was obtained. Our future work is to explore intra prediction and depth modeling modes in 3D-HEVC extension and apply the aforementioned algorithm to further effectively reduce the encoder computation complexity.

REFERENCES

[1] Gary J Sullivan, Jens Ohm, Woo-Jin Han, and Thomas Wiegand, "Overview of the high efficiency video coding (hevc) standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, 2012.

[2] Jani Lainema, Frank Bossen, Woo-Jin Han, Junghye Min, and Kemal Ugur, "Intra coding of the hevc standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1792–1801, 2012.

[3] Liang Zhao, Li Zhang, Siwei Ma, and Debin Zhao, "Fast mode decision algorithm for intra prediction in hevc," in *Visual Communications and Image Processing (VCIP), 2011 IEEE*. IEEE, 2011, pp. 1–4.

[4] Anil Fernando, Stewart T Worrall, and Erhan Ekmekcioð lu, *3DTV: Processing and Transmission of 3D Video Signals*, John Wiley & Sons, 2013.

[5] Yang Wang, Xiaopeng Fan, Liang Zhao, Siwei Ma, Debin Zhao, and Wen Gao, "A fast intra coding algorithm for hevc," in *Image Processing (ICIP), 2014 IEEE International Conference on*, Oct 2014, pp. 4117–4121.

[6] Wei Jiang, Hanjie Ma, and Yaowu Chen, "Gradient based fast mode decision algorithm for intra prediction in hevc," in *Consumer Electronics, Communications and Networks (CECNet), 2012 2nd International Conference on*. IEEE, 2012, pp. 1836–1840.

[7] Thaísa L da Silva, Luciano V Agostini, and Luis A da Silva Cruz, "Fast hevc intra prediction mode decision based on edge direction information," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. IEEE, 2012, pp. 1214–1218.

[8] Shunqing Yan, Liang Hong, Weifeng He, and Qin Wang, "Group-based fast mode decision algorithm for intra prediction in hevc," in *Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on*. IEEE, 2012, pp. 225–229.

[9] H Zhang and Z Ma, "Fast intra mode decision for high-efficiency video coding (hevc)," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 24, no. 4, pp. 660–668, April 2014.

[10] A. Dharmarajan and T. Velmurugan, "Applications of partition based clustering algorithms: A survey," in *Computational Intelligence and Computing Research (ICCIC), 2013 IEEE International Conference on*, Dec 2013, pp. 1–5.

[11] Frank Bossen, "Common test conditions and software reference configurations," *Joint Collaborative Team on Video Coding (JCT-VC), JCTVC-F900*, 2011.

[12] Gisle Bjontegaard, "Calculation of average psnr differences between rd-curves," *VCEG-M33*, April 2001.