Fast Coding Unit Size Selection for HEVC Inter Prediction

Thanuja Mallikarachchi, Anil Fernando and Hemantha Kodikara Arachchi
Centre for Vision Speech and Signal Processing, University of Surrey, United Kingdom
{d.mallikarachchi, w.fernando, h.kodikaraarachchi}@surrey.ac.uk

***Abstract*--Determining the best partitioning structure for a CTU is a time consuming operation for the HEVC encoder. This paper presents a fast CU size selection algorithm for HEVC using a CU classification technique. The proposed algorithm achieves an average of 67.83% encoding time efficiency improvement with a negligible rate-distortion loss.**

I. INTRODUCTION

Cisco's data traffic forecast statistics show that 80-90% of the global Internet traffic will be video data by 2017 and a significant proportion of the above percentage will be high definition content [1]. Hence, improved video compression techniques are required to manage this large volume of video data. With the intention of improving the coding efficiency to cater the upcoming video communication demands, High Efficiency Video Coding (HEVC) was standardized in early 2013 [2].

While inheriting most of the methodologies from its predecessors, HEVC introduces a number of new features. Similar to H.264/AVC, block based prediction and compression is the baseline for HEVC. However, a wider range of block sizes has been introduced [3]. In the *main profile*, a Coding Tree Unit (CTU) is partitioned into multiple Coding Units (CUs) of sizes ranging from 8×8 to 64×64. This flexible quadtree based partitioning structure is one of the main contributors for its improved coding efficiency [4]. Moreover, HEVC supports multiple PU/TU sizes and prediction modes that result in an enhanced coding efficiency compared to its predecessors [3].

Generally, HEVC compatible encoders employ rate-distortion (RD) optimization algorithms to determine the best prediction mode and the optimum CU size. With the increased number of CU/PU sizes and prediction modes, a significant proportion of the encoding time is spent on the RD optimization process. Numerous techniques have been reported in the recent literature to reduce its complexity. Approaches such as Motion Vector Merging [5], and that uses colour histogram features [6], achieve 34% and 45.33% encoding time reductions, respectively. However, [5] only determines the best PU mode while [6] evaluates RD cost for unnecessary depth levels before the decision is made, which incur an unnecessary execution time. The method proposed in [7] utilizes neighboring CUs for the current CU size decision, but requires intermediate frames with RD optimization to reduce the coding losses.

This work was supported by the ACTION-TV (http://www.action-tv.net/) project, which is funded under European Commission's 7th Framework Program (Grant number: 611761).

This paper introduces a fast CU size decision-taking algorithm for HEVC inter coding. The proposed algorithm utilizes motion and RD costs to classify CUs. The split probability for a given CU, which is subsequently used to make the split decision, is calculated using a probabilistic model. The proposed early termination prevents unnecessary CU evaluations resulting in an average encoding time saving of 67.83% while maintaining a similar RD performance.

The rest of the paper is organized as follows. Section II provides an illustrative overview of the proposed algorithm. Section III describes the experimental results and finally, Section IV concludes the paper.

II. PROPOSED METHOD

When considering the partitioning behavior of CUs with respect to inter prediction, it can be observed that blocks with homogeneous motion tend to utilize large CUs while blocks with complex motion utilize smaller CUs [5][6]. In this context, a content dependent CU classification technique is proposed. Obtaining the feature vector F for the classification is performed in two stages. First, *Inter N×N* mode is evaluated for the CU and a set of motion vectors are obtained for each constituent block.

$$MV = \{[mv_i, rPOC] \mid i = 0,1,2,3\} \tag{1}$$

where $rPOC$ is the Picture Order Count (POC) of the reference frame. The CU categorization based on MV set is depicted in Fig. 1. In this context, two motion vectors are considered to be similar when their horizontal and vertical components are equal and when they point to the same reference picture. The second feature is obtained from the RD cost computed for the *Inter N×N* mode [2]. The square root of the RD cost is quantized to the nearest integer and grouped into to bins to obtain $sqrdc_i$ where i=0, 5, 10, …, 200. Then the split decision, D_{split} of the CU is presented as a function of motion, RD cost and CU size.

$$D_{split} = f(Motion, RDCost, CUsize) \tag{2}$$

The decision making is performed through the dynamically formed probabilistic model based on the Bayes' theorem.

$$P(CU_S \mid F) = \frac{P(F \mid CU_S)P(CU_S)}{P(F)} \tag{3}$$

where $P(CU_s \mid F)$ is the probability of a CU is split, for a given feature vector F. From the conditional probability rules (3) is modified as,

$$P(CU_S \mid F) = \frac{P(CU_S \cap F)}{P(F)} \tag{4}$$

Fig. 1. Block classification based on motion homogeneity. Similar motion vectors are identified with the same index and color.

Each sequence undergoes an initial and intermediate training phases in which CU and PU mode decisions are taken through conventional RD optimization method. During these training phases, statistics on CU split decisions, motion category and RD costs are gathered. After the initial training phase, the subsequent frames are encoded with the CU split decisions derived from the probabilistic model based on the collected information. The equation (5) is utilized for the split probability calculation of the CU for a given feature vector *F*.

$$P\left(CU_S \mid F\right) = \frac{NCU_F^{split\ =1}}{NCU_F^{split\ =1} + NCU_F^{split\ =0}} \tag{5}$$

where *F* constitutes of motion category cat_i, $sqrdc_i$ and CU size and $NCU^{split=x}$ is the number of occurrences of CUs with x being the split decision. If $P(CU_s \mid F) > T$, the CU is split and if it is indeterminate due to insufficient data, the conventional RD optimization is used for taking the CU split decision, which initiate the intermediate training phase that will update the training data set. In this paper, the threshold T, is set to 0.6, which is determined through an empirical analysis of numerous video sequences.

III. Experimental Results

Simulations were conducted on a range of HD and CIF video sequences with diverse spatial and temporal characteristics. QP values were set to 22, 27, 32 and 37 and all the frames were encoded in *low delay P main* configuration in HM 12.0 encoder software [8]. The frame rates of HD and CIF sequences are 30 fps and 25 fps respectively. All the simulations were carried out on an Intel core i5 machine with 8GB RAM.

Table 1 summarizes the results with respect to average time saving (*ΔT*), Δ VQM [9], Δ PSNR and Δ Bit Rate against HM 12.0 [8]. The computational cost of the initial and intermediate training phases are also included in the time saving performance reported in this paper. Fig. 2 depicts RD and encoding time performance of the proposed method for a *CIF* sequence, with respect to HM12.0 [8], HM Fast methods [8] and two state-of-the-art methods described in [5] and [6]. Use of a low complex early termination algorithm that yield comparable split decisions similar to that of the RD optimization process makes the proposed method suitable for applications that require lower computational cost with a marginal coding efficiency loss.

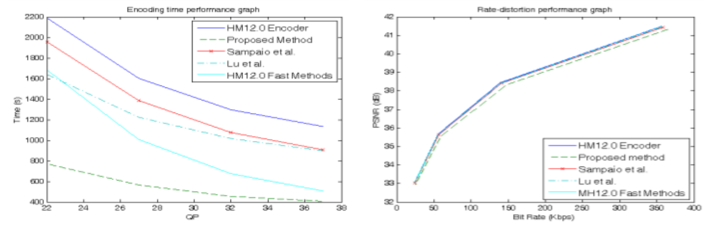

Fig. 2. Encoding time and RD performance for the ‘Container’ sequence.

IV. Conclusion

This paper proposes a fast CU size selection algorithm for HEVC inter prediction. A dynamic model for CU split decisions is formed with the motion characteristics and RD costs based on the initial training results. Subsequently, the model is updated based on the intermediate training results. Due to the early decision making made prior to the encoding of a CU, the proposed algorithm can provide an average time saving of 67.83% with a negligible impact on the PSNR, VQM and bit rate compared to the state-of-the-art algorithms.

TABLE I
Performance of the Proposed Algorithm

Sequence	*ΔT%*	Δ VQM	Δ PSNR	Δ Bit Rate%
Poznan Street 1088p	73.07	-0.05	0.06	3.55
City 720p	68.84	0.005	0.08	4.57
Kimono 1080p	69.92	0.003	0.05	3.85
Beergarden 1080p	69.60	0.002	0.07	5.44
Average	**70.35**	**-0.01**	**0.06**	**4.35**
Bridge-far CIF	71.47	-0.001	0.02	0.94
Highway CIF	60.27	0.005	0.11	7.00
Coastguard CIF	64.90	0.004	0.09	4.85
Container CIF	64.60	0.000	0.11	5.16
Average	**65.31**	**0.002**	**0.08**	**4.48**

References

[1] http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360_ns827_Networking_Solutions_white_Pap-er.html
[2] G. J. Sullivan, J. Ohm,W.-J.Han, and T.Wiegand, “Overview of the high efficiency video coding (HEVC) standard”, *IEEE Trans. Circuits Syst. Video Technol*., vol. 22, no. 12, pp. 1649–1668, Dec. 2012.
[3] I.-K. Kim, J. Min, T. Lee, W.-J. Han, and J.-H. Park, “Block Partitioning Structures in HEVC”, *IEEE Trans. Circuits Syst. Video Technol*., vol. 22, no. 12, pp. 1697 – 1706, Dec. 2012.
[4] J. Ohm, G.J. Sullivan, H. Schwarz, T. K. Thiow, and T. Wiegand, “Comparison of the Coding Efficiency of Video Coding – Including High Efficiency Video Coding (HEVC)”, *IEEE Transactions on Circuits and Systems for Video Technology*., vol. 22, pp. 1669 – 1684, Dec. 2012.
[5] F. Sampaio, S. Bampi, M. Grellert, L. Agostini, and J. Mattos, “Motion Vectors Merging: Low Complexity Prediction Unit Decision Heuristic for the Inter-prediction of HEVC Encoders”, *IEEE International Conference on Multimedia and Expo (ICME)*., Jul. 2012, pp. 657 – 662.
[6] J. Lu, F. Liang, L. Xie, and Y. Luo, “A fast block partition algorithm for HEVC”, *International Conference on Information, Communications and Signal Processing (ICICS)*, Dec. 2013, pp. 1 – 5.
[7] W.-J. Hsu and H.-M. Hang, "Fast Coding Unit Decision Algorithm for HEVC", *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Nov. 2013, pp. 1-5.
[8] https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-12.0/.
[9] M.H. Pinson and S. Wolf, “A new standardized method for objectively measuring video quality”, *IEEE Transactions on Broadcasting*., vol. 50, pp. 312-322, Sept. 2004.