International Journal of Engineering & Technology, 7 (4) (2018) 4499-4502

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 19450
Research paper

HEVC intra prediction analysis for optimization

V. Sulochana [1] *, B. Shanthini [2], G. B. Praveen [3]

[1] *Research Scholar - Department of CSE, St. Peter's University, Chennai, India*
[2] *HOD- Department of IT, St. Peter's College of Engineering and Technology, Chennai, India*
[3] *Technical Lead-Path Partner Technology PVT Ltd, Bangalore, India*
**Corresponding author E-mail: sulo62002@yahoo.com*

Abstract

Increasing demand for video data has driven new video compression standard HEVC successor of H.264/AVC which is intended to achieve 50% gain in compression efficiency. Compared to H.264 which had only 9 intra prediction mode, more angular intra prediction mode directions are introduced in this standard. Increase in the performance proportionally increases the complexity of encoder and hence a fast method is required to choose the optimal direction. In this work, we have studied existing literatures that are envisioned to reduce intra coding time and briefly presented the comparison for future enhancement.

Keywords: *RDO; DCT; SATD; RMD; MPM*

1. Introduction

HEVC provides 50% bit rate reduction with same visual quality compared to H.264. HEVC is designed with advance flexible coding structure, where each pictures are divided in to tiles or slices for further processing. Coding Tree unit (CTU) is the elementary processing unit of HEVC, which is similar to MacroBlock (MB) in H.264. Dedicated syntax elements for both Luma and chroma block are part of CTU which forms the Coding Tree Block (CTB). The basic units of HEVC are Coding Unit (CU), Prediction Unit (PU) and Transform Unit (TU) [1]. All these units contain individual syntax elements for Luma & Chroma. Recursive splitting of CU forms an optimal quad-tree structure and are traversed in Z-scan order. In HEVC, possible CU sizes are 8×8, 16×16, 32×32, 64×64.

Conventionally, complex texture regions are split into smaller CU's and homogeneous texture regions are coded as larger CU for maintaining better compression efficiency. Each CU is further sub-divided into possible PU sizes as defined in standard. As defined above, estimation of all possible CU sizes is highly time consuming which degrades real-time capability of encoder. Estimation of 35 directions for each possible intra PU is additional overload [2].

The aim of this paper is to intensively survey the various approaches that are existing to reduce computation time of intra PU's. Section II describes the intra-prediction framework and section III describes the various intra-prediction mode decision techniques.

2. Intra prediction framework in HEVC

2.1. Prediction overview

Aim of the prediction is to find data redundancy which in turn improves the compression efficiency. Prediction of current block using pixels from same frame is termed as intra prediction and if prediction is done using pixels from nearby frames, it is termed as inter prediction. In HEVC symmetric prediction block sizes are used in both inter and intra prediction whereas asymmetric block sizes are used only for inter prediction [3]. In intra prediction, every pixel predicted is weighted combination of neighbouring pixels depending on intra-prediction angle used [4]. Block estimation can be done by set of bounding pixels.

2.2. Intra prediction modes

HEVC provides 35 different modes for prediction:

a) DC mode (mode 0)
b) Planar mode (mode 1)
c) 33 angular modes (mode 2-34)

Conventionally, if region to be predicted is homogeneous (or) smooth with less texture, DC/planar modes are best suited most of the times and in case textures are present then angular modes are well suited. Basic steps in HEVC Intra prediction are construction of reference samples and block prediction using constructed reference samples [5].

2.3. Reference sample

Pixels from neighbouring reconstructed blocks are used as reference samples. Samples are referred from left block [6] (marked in blue), top block (marked in green), top-left (marked in red), top-right and bottom left (marked in brown) as shown in Fig 1.If particular neighbour block is not available, samples for corresponding position is attained from nearest available samples [7].

2.3.1. Filtering reference samples

The reference samples used by HEVC intra prediction are conditionally filtered by a smoothing filter. For better prediction, spurious noise and unwanted directional edges are removed during filtering process. Depending upon the prediction, block size and prediction mode filtering option is enabled. In general three-tap filters are used [8]. In case of DC mode prediction and 4x4 block size, filter is turned off. For 8x8 block, filter is applied only for diagonal direction angular modes 2, 18 and 34. For 16x16 block size, filters are not applied for angular modes 9, 10, 11, 25, 26 and 27. For 32x32 block size, angular modes 10 and 26 are excluded from filtering process.

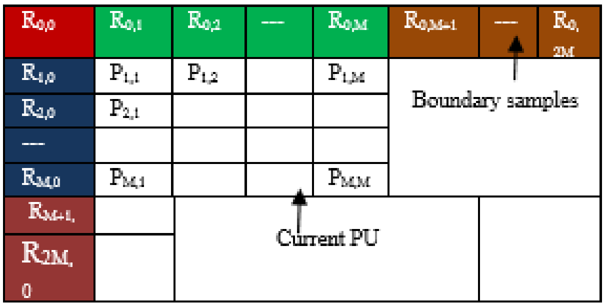

Fig. 1: Reference Sample.

2.4. Angular modes

Intra angular mode predictions are based on the specific direction. Angular modes are indexed from 2 to 34 as shown in Fig 2. Directions are from near horizontal through near-diagonal to near-vertical. Horizontal direction is indexed as Mode-10 and vertical direction as Mode-26.
Chroma prediction can be explicitly signalled or can be inferred from mode for luma prediction. INTRA_DERIVED mode is signalled when same luma prediction is used. In Chroma prediction, possible five modes are planar, DC, vertical (26), horizontal (10) and derived.

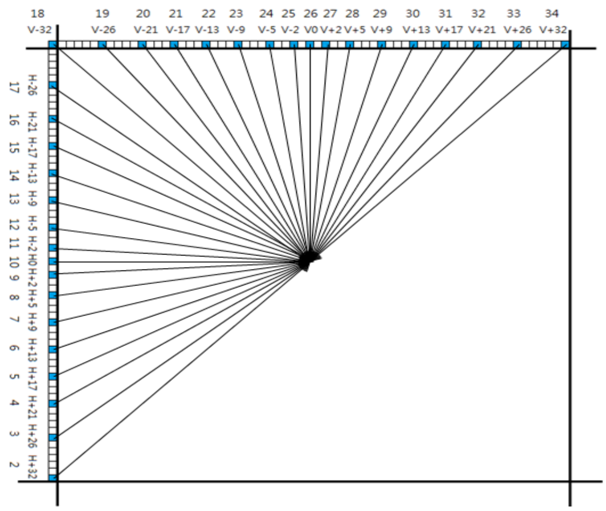

Fig. 2: Angular Modes.

2.5. Intra sample prediction

To predict various kinds of contents, HEVC supports a range of sample prediction methods. Rate Distortion (RD) calculation is the rate at which the amount of distortion against the amount of data required to encode the video. Within intra prediction, one of the time consuming sub module is to predict the best mode using RD calculation [9].

2.5.1. Reducing number of modes

In HEVC, since the number of angular modes is 33 which are high in numbers to perform full Rate Distortion Optimization (RDO) process, modes are reduced to N candidates. At first, Rough Mode Decision (RMD) [10] process shortlisted the 33 angular mode into N modes using Sum of Absolute Sum of Absolute Transform Differences (SATD) which is computed based on Hadamard transform and not Discrete Cosine Transform (DCT). Hadamard needs only integer addition operation hence it reduces the computation time than DCT.
Next step is to obtain Most Probable Mode (MPM) [11] from N modes through previously decoded two neighbouring PU's, if the MPM modes are not available in the candidate list it will be added. Now simplified Rate Distortion (RD) is obtained for candidate list, from RD cost best mode is selected and full RD is calculated.
In intra prediction, optimization can be majorly done in two units CU and PU. Fast coding unit size selection in CU and Fast mode decision in PU are the major areas to explore.

2.5.2. Fast coding unit size selection

Majorly early termination of tree splits on the fly and pre-decision of tree splits are the two major approaches for fast CU size selection. In case of early termination, RD cost is calculated based on the statistics and threshold value is used to terminate the splitting process.
Other possible approaches are using weighted SVM and Bayesian decision rule to implement CU splitting, Bayesian theory to reduce the TU size, machine learning approach with set of decision trees, by using the directional features like mean, variance and mean absolute difference in vertical and horizontal directions to decide on CU early terminations are adopted.

2.5.3. Fast mode decision

In case of fast mode decision, pixel gradient direction is computed using filters and applied to find out dense occurrence of gradient. This type of approach is extensively used for best angular mode predictions to reduce the number of modes during RMD and MPM. To reduce the number of modes other possible approaches used are edge detection based methods, using output of filters in decision making, directional mask based methods, gradient intensity based methods and statistical correlations techniques

3. Intra mode prediction techniques

In this section, we have discussed about existing approaches given by various authors on reducing computation time of intra prediction mode.
In intraprediction, dominant directions are found with different angular step size. Angular step size is calculated till the best mode is predicted. To find the directions each PU is converted to transform domain, based on directions classified intra-mode groups is offered as candidate group.
Based on co-efficient distribution of DCT, modes are grouped and if the co-efficient value is greater than threshold value then it's not flat region. If it contains more active regions edges may be available. Edge information is used to decide the direction based on the proportion of intensity. Intra mode grouping are based on major directionality and the groups are categorised under horizontal, vertical, diagonal_45º, diagonal_135º, non-directional and flat. Intra mode group needs to be updated iteratively since it uses same intra-mode group until next PU starts encoding. With this algorithm, time savings is more than 25% of average but with trivial visual quality loss [12].
To reduce the number of candidate modes, SATD cost is arranged in ascending order of i to i+1. Hence the probability of ith mode to be selected is more.

$$SatdCost\ i+1 - SatdCosti \leq \alpha j.\ SatdCost\ i+1 + SatdCosti \quad (1)$$

If ith mode is selected with equation 1 also satisfied, then (i+1) mode is selected else mode is rejected. If α increases, then more candidates are included in the list. Varying factor (α) for all the current blocks are calculated by linear weightage of neighbouring blocks. This method can save 32% of encoding time by only 0.9% BD-BR increase [13].
The modes are grouped into statistically homogeneous group from MPM and RMD. Clustering based k-medoids algorithm is used to

group the modes within each cluster to find the best mode. In k-medoids, instead of mean value of object it takes object distribution as a center for cluster. The non- medoids objects are identified based on the distance from the center will be assigned to near neighbouring clusters [14].

Gradient is computed using Sobel operator and using gradient direction with maximum variation in the pixels, where the edges perpendicular to this direction are used to find the best mode for intra-prediction. First step is to find the edges to get three related modes, then DC, planar and best modes of five neighbouring block modes are added to the list. From SATD cost of the list N modes are selected and binary classification is applied that is powerful mode with low cost and least mode with high cost are separated based on the adaptivity of block size and difference between the maximum and minimum cost values. Suggestion given in this paper is classification can be done with various parameters like quantization parameter, or absolute minimum and maximum cost values or any other criterion. This approach achieves 35.6% time reduction on average, with minimum degradation in coding efficiency [15].

Gradient vectors for horizontal and vertical directions Gx and Gy are computed by sobel operator. By using the horizontal and vertical gradient vectors angle and amplitude are defined. For RMD a mode with little higher amplitudes are selected. Conclusion in this paper is if average of Gx / Gy is greater than 1 best modes are from vertical and if Gy / Gx is greater than 1 then the best predicted modes are from horizontal and if average of horizontal and vertical is very small then probably the modes will be DC and planar. In this approach encoding time saving is about 54% on average with 0.7% BD-rate increase [16].

To reduce the number of candidate modes, the threshold value is obtained by the cost difference between the first and last selected mode during RMD. By the same way MPM is also reduced using cost threshold. This method achieves 41.8% reduction in time with increase in bit rate 0.07% and quality loss over 0.077 dB [17].

In intra prediction, a linear mapping by trained sets between known and reference pixel are iteratively done by k-means clustering to get the best prediction. A linear mapping is dynamically changed during each iterative process to get the best results. Mapping can be done by grouping similar blocks and it forms a cluster. This approach is quiet complex, for 'N' pixel to predict from 'M' reference pixel it requires 'M * N' multiplication for each group, but it is more consistent with quality predictions and with additional hardware support performance can be improved dramatically [18].

To simplify RD cost, Rough Mode Cost (RMC) is calculated for all intra modes. CU partitioning, PU Intra mode Decision and TU depth decision is carried out based on RMC. CU Early termination or Early splitting can be decided by RMC value. In this approach RMD for CU is obtained before processing and hence pre-calculated RMC for intra mode decision can be done in advance. For PU mode decision, high RMC value has less probability to be a best mode hence ratio between the least RMC mode with respect to all modes are used to reduce the candidate mode. With TU partition, least RMC mode is used to predict the TU partition of other modes. This algorithm reduces encoding time by an average of 51% with 0.69% increase in BD rate [19].

RMD is applied only to 19 modes out of 35. Rough encoding cost of two neighbouring modes which are already defined for previous RMD is used to exclude the modes during RMD. First step is to calculate the cost for half of the angular direction modes that is even mode along with DC and planar are obtained by Hadamard transform. Compute the remaining mode where the cost is based on average of two neighbouring modes, at the stage of MPM further modes are reduced based on the correlation among the modes. To further reduce the best mode, considerably the vertical and horizontal modes have least correlation, hence any one can be picked up for RDO process. This method achieves reduction in encoding time by 14.4% with intra-main and 7.2% reduction with all intra-main10 configurations [20].

Texture orientation detection algorithm, in this directional variance are computed in specific spatial direction with rational slopes in lattice and dense orientation texture is chosen from the lowest Mean variance. Variance is used to reduce the candidate set mode at RDO stage. Only 12 co-lines with relational slope are considered instead of 33 directional modes. In the 12, four with same slope horizontal, vertical and two with diagonal orientation are selected, then 8 slopes which are close to a mode with orientation point between two reference samples with rational slope of +-1/2, +-1/4,+-2 and +-4 are selected. Lower variance is considered to be a dominant texture orientation to reduce the modes. This method achieves 30% reduction in computational complexity with increase in bit rate 0.4% and quality loss over 0.03 dB [21].

In HEVC, there are 4N+1-reference samples are used for intra prediction. In the reference sample, if there is continuity in the structure then based on the direction of the structure angular modes are selected. First step is to find out whether the reference sample is structured or homogeneous. For the sample, reference 4N+1 variance is calculated. If the variance is less than threshold value samples are considered as homogeneous or else its structured. If the reference sample is homogeneous, angular modes are skipped. Exclusion of modes reduce the time more than 20% by bitrate increase of only 0.7% [22].

The Distortion variance, which is the difference between the current, and reference macro blocks are used to predict the mode. Instead of calculating SAD, ASD (Absolute Sum of Difference) that is the difference between the luminance pixel value of adjacent and current MB between two consecutive frames are calculated [23]. Instead of sobel, prewitt operator is used to generate the gradient. Generally sobel operator is one of the best edge detection operator, but prewitt offers simpler coefficients than sobel. Prewitt operator uses +1 and -1 as coefficient that requires simple addition and subtraction compared to Sobel which uses +2 and -2. A gradient candidate list is calculated from gradient set being the appearance of number of modes in PU. A candidate list after RMD is compared with gradient candidate list to find the modes based on certain conditions. The proposed algorithm achieves a time saving of 42.8% with an average increase in BD-rate of just 1.1% [24].

To reduce the number of candidate mode after RMD, the relationship between RMD values of candidate list and the position of the best mode in reordered candidate list is defined to skip the modes further. To define the criteria the difference between the first and last reordered list are evaluated. If the difference is greater than threshold corresponding modes are skipped for full RD. It shows better performance when threshold is 0.35 and quality loss over 0.0051 dB [25].

4. Conclusion

In this work, existing literatures are intrinsically surveyed and analogized to reduce intra coding time. Methods are compared on the scale of compression efficiency and encoding time reduction. All the above methods provide substantial reduction in encoding time and at the same time some of the approaches are resulting in trivial coding loss. To improve the performance in future, surveyed methods can be further enhanced and optimized. This work can serve as a base for exploring further approaches in reducing intra coding time without much degradation in compression efficiency.

References

[1] Andrzej Abramowski, "A survey over possible intra prediction optimizations in the H.265/HEVC encoder", Proc. of SPIE, Vol: 10031 1003153-13, December 2016.

[2] Frank Bossen, Benjamin Bross, Karsten Suhring, and David Flynn, "HEVC Complexity and Implementation Analysis" IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, pp. 1648–1667, Dec. 2012 https://doi.org/10.1109/TCSVT.2012.2221255.

[3] Jing He, Wenkao Yang, and Jing Wang, "Fast HEVC Coding Unit Decision Based On BP-Neural Network" International Journal of

Grid Distribution Computing., vol. 8, no.4, pp. 289–300, 2015 https://doi.org/10.14257/ijgdc.2015.8.4.28.

[4] G. J. Sullivan, J.-R. Ohm, W.-J. Han, Heiko Schwarz, Thiow Keng Tan and T. Wiegand, “Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding,” IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 12, pp. 1669, Dec. 2012 https://doi.org/10.1109/TCSVT.2012.2221192.

[5] Kevin REUZE, Pierrick PHILIPPE, Oli vier DEFORGES, Wassim Hamidouche, “Enhanced Intra Prediction Modes Signalling in HEVC” IEEE, 978-1-5090-5966-9/16,2016.

[6] Heming SUN, Dajiang ZHOU, Peilin LIU, and Satoshi GOTO, “Fast Prediction Unit Selection and Mode Selection for HEVC IntraPrediction,” IEICE Trans. Fundamentals., vol. E97-A no. 2, Feb. 2014.

[7] Manjanaik.N, Manjunath.R, “Selection Of Intra Prediction Modes For Intra Frame Coding In Advanced Video Coding Standard”,IJRET, ISSN:2319-1163., vol.2,no.12,Dec 2013.

[8] G. J. Sullivan, J.-R. 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. 1648–1667, Dec. 2012 https://doi.org/10.1109/TCSVT.2012.2221191.

[9] Nan Hu, En-hui Yang, “Fast Mode Selection for Hevc Intra Frame Coding with Entropy Coding Refinement Based on Transparent Composite Model,” IEEE Trans. Circuits Syst. Video Technol., pp. 1051–8215., Dec. 2015.

[10] Hao Liu, YuexinJie, “Fast HEVC Intramode Decision Based on Hybrid Cost Ranking”, Advances in Multimedia Volume, Hindawi Publishing Corporation, Article ID: 5164025, 2016.

[11] Jakub Siast, Jakub Stankowski, Marek, “Jakub Siast, Jakub Stankowski, Marek, “Intl Journal of Elctronics and Telecommunications”, vol. 62, no: 2, pp. 147–151, Mar 2016.

[12] Miso Park, Jechang Jeong, “Fast HEVC IntraPrediction Algorithm with Enhanced Intro Mode Grouping Based on Edge Detection in Trasnform Domian”, Journal of Advances in Computer Networks, vol.: 3, no: 2 , June 2015.

[13] Fu Li & Dandan Jiao & Guangming Shi, “An AR based fast mode decision for H.265/HEVC intra coding”, Springer, Multimed Tools Appl, DOI 10.1007/s11042-016-3737-z, July 2016.

[14] Sami Jaballah, Kais Rouis, Jamel Belhadj Tahar, “Clustering-Based Fast Intraprediction Mode Algorithm For Hevc”, IEEE(EUSIPCO), 978 8-0-9928626-3-3, 2015

[15] Mohammadreza Jamali1, Stéphane Coulombe1, François Caron, “Fast HEVC Intra Mode Decision Based on Edge Detection and SATD Costs Classification”,IEEE(DCC),2015

[16] Tao Zhang, Ming-Ting Sun, Debin Zhao, Wen Gao, ” Fast Intra Mode and CU Size Decision for HEVC”, IEEE, vol.: 1051-8215, 2015

[17] Kanayah Saurty,Pierre Clarel Catherine, K.M.S.Soyjaudah, “An Improved Reduced Candidate Modes Approach for HEVC Intra Prediction”, IEEE, ISBN:978-1-4673-7606-8/15, 2015

[18] Carlo Noel Ochotorena, Yukihiko Yamashita, “Regression-based Intra-prediction for Image and Video Coding”,IEEE, arXiv:1605.03754v1,2016.

[19] Zong-Yi Chen, Pao-Chi Chang, “Rough Mode cost-based fast intra coding for high-efficiency video coding”, Elsevier,J.Vis.Commun.ImageR, vol.:43,2017.

[20] Mohammadreza Ramezanpour Fini, Farzad Zargari, “Two stage fast mode decision algorithm for intra prediction in HEVC”, Springer, Multimed Tools Appl, DOI 10.1007/s11042-015-2675-5, 2015.

[21] Damian Ruiz, Gerardo Fernandez-Escribano, Jose Luiz Martinez, Pedro Cuence, “Fast Intra mode decision algorithm based on texture orientation detection in HEVC”, Elsevier, Signal Processing: Image Communication, vol.:44,2016.

[22] Andreas Heindel and Andr´e Kaup, “Fast Exclusion of Angular Intra Prediction Modes in HEVC Using Reference Sample Variance”, IEEE, ISBN: 978-1-4799-5341-7/16, 2016.

[23] L. Balaji, V. Logesh Raj, R. Raghavendran, S. Mathan Raj, G. Akash, “Fast Intra Mode Prediction for HEVC Based on Distortion Variance”,IEEE, International Conference on Circuit, Power and Computing Technologies,ISBN: 978-1-5090-1277-0/16,2016.

[24] Anis BenHajyoussef, Tahar Ezzedine, Ammar Bouallègue, “Gradient-Based Pre-Processing For Intra Prediction in High Efficiency Video Coding”, Springer, EURASIP Journal on Image and Video Processing, https://doi.org/10.1186/s13640-016-0159-9.

[25] Do-Kyung Lee, Miso Park and Je-Chang Jeong, “Fast Intra Coding by using RD Cost Candidate Elimination for High Efficiency Video Coding”, WCECS, ISBN: 978-988-19252-0-6, 2014.