iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** III **Month of publication:** March 2015

DOI:

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

A Survey: Fast Inter CU selection algorithms for HEVC

R. S. Hingane[1], S. S. Agrawal[2]
Department of E&TC University of Pune,India.

***Abstract—* A hybrid quad tree structure in which current CU recursively divided into four equal sized CU's, is adopted by HEVC. At each decomposition level (depth), the RD cost for each of the following modes, SKIP mode, merge mode, inter 2N × 2N, inter 2N × N, inter N × 2N, inter 2N × nU, inter 2N × nD, inter nL × 2N, inter nR × 2N, inter N × N (only available for the smallest CU), intra 2N × 2N, and intra N × N (only available for the smallest CU) in inter-frame prediction is calculated. Among the all possible depth levels and prediction modes, the one which having least RD cost is selected. These inter CU selection process poses high computational complexity which is main obstacle in implementing HEVC for real time applications. Recently huge research has carried out on reducing computational complexity of HEVC encoders. This paper presents different methods for fast inter CU selection algorithms which reduces the encoding time of HEVC coding process. A survey of these various algorithms and an evaluation of their pros and cons may provide valuable leads for the improvement in HEVC inter-prediction.**
***Keywords—* H.264/AVC;H.265/HEVC;CU; inter prediction; RD cost.**

I. INTRODUCTION

THE Joint Collaborative Team on Video Coding (JCT-VC) formed collaboratively by ISO-IEC/MPEG and ITU-T/VCEG, has developed new video coding standard, High Efficiency Video Coding (HEVC)/H.265. HEVC adopted a quad tree coding structure with Coding Unit (CU) as a basic coding element. Unlike macroblock (MB) in AVC/H.264 standard, CU size is flexible (from 64 X 64 to 8 X 8 luma samples). A CU is recursively divided into four equally sized CU's which enables content adaptive coding in HEVC. CU splitting continues until it could not be divided further into smaller CU's. The Prediction Unit (PU) carries the information related to prediction process. The size of PU is not restricted to be square to occupy boundaries of the objects in the pictures.

Fig. 1 shows CU partitioning up to depth level 3. The dark black lines shows CU at level 0, blue lines split CU into depth level 1 and so on. CU is encoded by deciding its optimal prediction mode out of 10. At each depth level, RD cost for each of the prediction modes is calculated. The prediction mode having least RD cost value is selected. After deciding on RD costs for all CU's at depth level, say X, all four cost values are added and this addition is compared with RD cost of corresponding CU in previous depth level X-1. If addition value is less than cost value at depth X-1, then CU splitting is continued (up to level 3), else splitting is stopped at depth level X-1. For deciding on CU splitting up to particular depth level and PU partitioning mode, HEVC encoder needs to calculate large number of RD costs. The computing process of RD cost contains a series of operations as prediction, transform/ inverse transform, quantization/ dequantization, and entropy coding for each possible PU mode which makes major contribution in computational complexity.

Recently heavy research has been carried on reducing the computational complexity of HEVC encoding process. This paper presents survey on different methods for inter CU selection algorithms. The paper proceeds as; Section II Overview of HEVC Encoding, Section III gives Literature Review of different methods for fast inter CU selection, Section IV gives summary of survey and conclusions are given in Section V.

II. OVERVIEW OF HEVC CODING STRUCTURE

A. HEVC coding structure

The various terminologies involved in hybrid video coding for HEVC are highlighted as follows.

1) *Coding tree units (CTU) and coding tree block (CTB):* In the H.264/AVC standard, the macroblock (MB) was the basic coding unit containing a 16×16 block of luma samples and two corresponding 8×8 blocks of chroma samples in the color sampling format of 4:2:0. Whereas in HEVC, the coding tree unit (CTU), supports size selected by the encoder varying from 16 X 16 to 64 X 64. The CTU consists of a one luma CTB and the corresponding two chroma CTBs and syntax elements. HEVC supports a quad tree partitioning of the CTBs into smaller blocks.
2) *Coding units (CUs) and coding blocks (CBs):* A CTB may split into four coding units (CU). CU is a basic coding element in HEVC. Each CU has one luma coding block (CB), two corresponding chroma CBs and syntax elements and has an

associated partitioning into prediction units (PUs) and a tree of transform units (TUs).

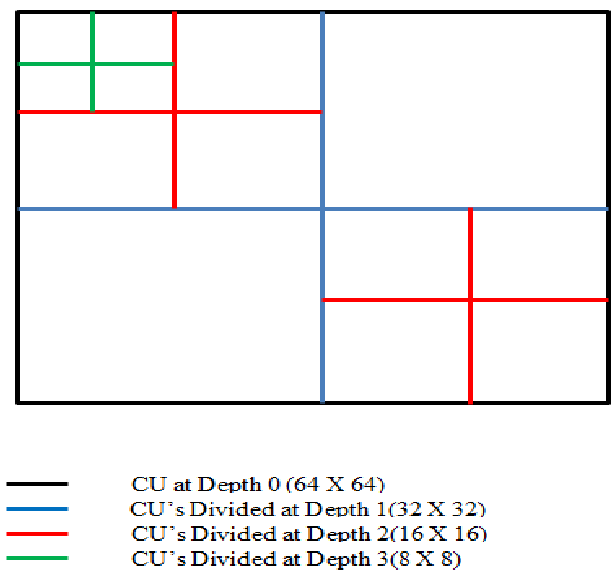


Fig. 1 CU Splitting at different depth levels.

3) *Prediction units (PUs) and prediction blocks (PBs):* Inter or intra prediction decision is made at the CU level using prediction units (PU). Depending on the basic prediction-type decision, the luma and chroma CBs can then be further split in size and predicted from luma and chroma prediction blocks (PBs). HEVC supports variable PB sizes from 64×64 down to 4×4 samples.

4) *TUs and transform blocks*: A block-based discrete cosine transform (DCT) is used to transform predicted residual to form transform unit (TU). The size of TU is from 32x32 to 4x4. Discrete sine transform (DST) is specified as alternative transform for the 4x4 transform of luma intra residual.

This classification of the block structure into three different formats (CU, PU, and TU) provides scope for optimization of each format according to its role.

In a case where multiple-reference pictures are there, a particular set of previously decoded pictures needs to be present in the decoded picture buffer (DPB) for the decoding of the remainder of the pictures [ref]. To identify such pictures, a list of picture order count (POC) identifiers is transmitted in header of each slice. There are two such lists called reference picture list 0 (L0) and list 1 (L1). To identify a particular picture in one of these lists, an index called a reference picture index is used. A picture is selected depending upon prediction mode i.e. if uniprediction, a picture is selected from either of these lists and for biprediction, two pictures are selected, one from each list. If a list contains only one picture, the reference picture index implicitly has the default value 0 [1].

B. Analysis of Inter-prediction in HEVC

The process of inter prediction consists of merge estimation and conventional Motion Estimation (ME) to obtain motion vector (MV), reference frame index, and inter prediction mode information. Fig. 2 a) shows flowchart for the process of inters prediction in HEVC. HEVC has adopted a new coding tool, Merge estimation. Merge estimation derives the motion information from spatially or temporally neighbouring blocks [3]. ME is carried out for the three inter prediction modes such as uni-directional prediction in List 0 (Uni-L0), uni-directional prediction in List 1 (Uni-L1), and Bi-directional prediction (Bi). The best inter prediction mode of current PU is finally decided among Uni-L0, Uni-L1, Biprediction modes by Rate Distortion Optimization (RDO) [2].

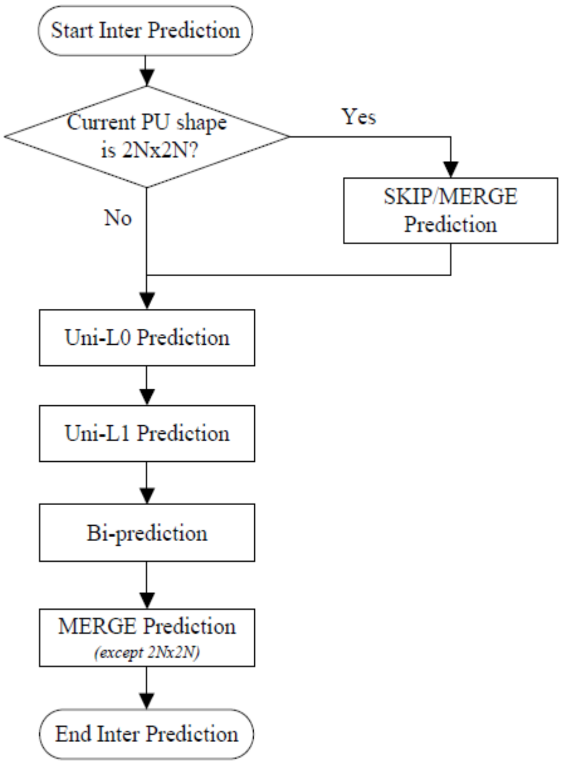
As shown in Fig. 2 b), out of the total encoder complexity, inter prediction contributes about 69%, transformation and quantization makes contribution to 25% and intra prediction and remaining encoding operations contributes 3% each [3]. Hence any efforts regarding reduction in inter prediction complexity will take real time implementation of HEVC one step forward.

III. SURVEY OF INTER CU SELECTION ALGORITHMS

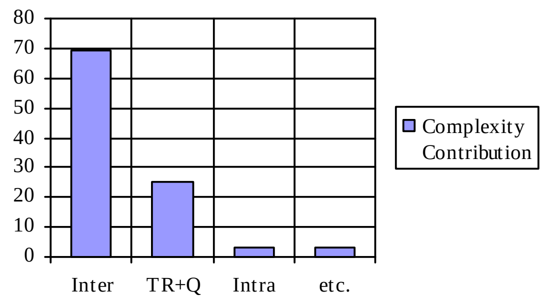
Volume 3 Issue III, March 2015
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Recently, studies on reducing the computational complexity of HEVC encoders are reported.

A. In the method proposed in [4], at frame level, a fact that for several consecutive frames, some encoding features like Quantization Parameter (QP), motion speed and resolution remains constant is used. Hence by exploiting mode



a)



b)

Fig. 2 a) Inter prediction process, b) Analysis of complexity contribution.

correlation between consecutive frames it skips some specific depth levels which were rarely used in previous frames. At CU level the method relies on the fact that the features like motion and texture detail of one particular portion remains the same. By exploiting the neighbor CU and co-located CU information, it pre-determines the candidate CU depth. This algorithm shows averagely 45% total encoding time reduction and 0.19% bit rate increment.

B. The method in [5] adaptively determines the CU depth range by using treeblock property that small depth levels are always selected at treeblocks in the homogeneous region, and large depth levels are selected at treeblocks with active motion or rich texture, so that motion estimation can be skipped.

C. Algorithm that is easy to implement, proposed in [6] detects SKIP mode in one CU-level based on the differential motion vector (DMV) and coded block flag (CBF). It shows that CU searches inter 2Nx2N mode before checking the SKIP mode. Then method checks its DMV and CBF. If the DMV and CBF of the inter 2Nx2N mode are equal to (0, 0) and zero respectively, where these two conditions together are called as early SKIP conditions, the best mode of CU is determined as the SKIP mode i.e. the remaining PU modes are not searched for inter mode decision. The method shows the encoding complexity reduced by up to 34.55% in random access (RA) configuration and 36.48% in low delay (LD) configuration of HM 4.0 software with only a bit of rate increment of 0.4% in RA and 0.5% in LD configurations respectively.

D. The method concentrating on rate distortion optimization (RDO) in HEVC is proposed in [7]. Top Skip and Early Termination are used to reduce number of RD cost estimations. In Top Skip, the starting depth (or the larger depths like 64 X 64 and 32 X 32) is skipped if there is high correlation between the minimum depth of the current CTB and the one of the co-located CTB in the previous frame. Early termination is complementary process to Top Skip and stops CU splitting process if the best RD cost for current CU is lower than the threshold. The threshold is adaptively calculated by exploiting both spatial and temporal features of frames. The method reduces the encoding time, notably up to 45% regarding the high complexity HEVC reference software codec, with a negligible quality loss never larger than 0.1 dB.

E. The Bayesian decision rule based algorithm is proposed in [8]. It collects relevant and easily computable features to help decisions on CU splitting. The features are collected using mutual information between classifiers and feature vectors. A Bayesian decision rule is defined to reduce RD loss, based on which the encoder reduces 41.4% encoding time on average (up to 66%) at the cost of 1.88% loss on RD performance.

F. Another spatio-temporal correlation based method is proposed in [9] in which, depending upon correlation index of current CTU with its spatio-temporal neighbor, CTU is classified into three categories namely High Similarity, Medium Similarity and Low Similarity. CTU's in High Similarity category will have smaller depth search range, in Medium Similarity will have medium while Low Similarity will have go through whole depth range i.e. (from 0 to 3). Algorithm provides avg. 25% time savings with 0.16% bit rate increment.

G. Modeling based CU selection method is proposed in [10]. It uses the statistical observation that CU's with same pyramid variance of the absolute difference (PVAD) tend to get encoded at same depth. The CU selection process is modeled as Markov Random Field (MRF) inference problem optimized by Graph cut algorithm and termination of CU splitting is decided by a RD cost based maximum a posteriori (MAP) approach. Algorithm gives avg. 53% of time savings and bit rate increment of 1.32%.

H. Binary classification modeled CU selection method is proposed in [11]. An offline Support Vector Machine (SVM) model optimized by using weights on training samples is used to classify the CU into categories as Split or Unsplit. The weights are generated as the RD cost difference due to misclassification. SVM is applied on feature vector to make CU splitting termination decision. Method shows 44.7% complexity reduction and 1.35% bit rate increment in RA configuration and 41.9% time savings with 1.66% bit rate increment in LD configuration.

I. Bayes rule based early CU splitting and pruning decision method is proposed in [12]. In early split test, Bayes rule based on low complexity RD cost decides whether current CU could be splitted or not. If not then CU early pruning test is carried out. Pruning decision is taken based on full RD cost of CU. Features for both the tests are periodically updated online. The method gives about 50% encoding time reduction with only 0.6% increment of bit rate.

J. Spatio-temporal correlation based complexity control method is proposed in [13]. Here a predefined target for complexity of HEVC encoder is achieved at the cost of small reduction in RD performance. The method dynamically adjusts CU structure to meet the complexity target. CU structure adjustment is based on Spatio-temporal correlation of current CU. The method shows 40% reduction in complexity with 1.42% bit rate increment.

K. A hierarchical complexity allocation at frame level using linear programming is proposed in [14] for fast CU size decision. A mode mapping approach is used along with allocated complexity factor for given CU, to take decision on that CU.

L. Method in [15] adjusts the weighting factor of Fast ENcoder (FEN) based on probability of CU getting encoded in SKIP mode. Probability is calculated using mode decisions taken on previously encoded spatial neighborhood CU's. Adaptive adjustment of weighting factor reduces loss of RD performance. It is update for FEN having same encoding time and improved RD performance.

M. Pyramid Motion Divergence (PMD) based fast inter CU selection method is proposed in [16]. It uses the statistical property that CU's having similar PMD value will have same splitting mode. Optical flow is evaluated to find MV's so that PMD can be estimated. But optical flow estimation causes computational burden of 9%. Still method provides avg. 41% time savings with avg. 1.9% bit rate increment.

N. The extension to method in [10] is proposed in [17]. Here instead of using PVAD, only VAD is used because of the observation that VAD is proportional to the RD cost. Remaining is similar to [10]. It provides 53% time savings with bit rate increment of 2.74%.

O. Method utilizing the byproducts of HEVC encoding process so that no extra computational burden on encoder is proposed in [18]. It makes use of sample-adaptive-offset (SAO) parameters as the spatial encoding parameter to estimate the texture complexity while, the motion vectors, TU size and coded block flag (CBF) information are used as the temporal encoding parameters to estimate the temporal complexity. Method provides 49.6% complexity reduction and 1.4% bit rate increment in RA configuration and 42.7% time savings with 1.0% bit rate increment in LD configuration.

TABLE I

SUMMARY OF INTER CU SELECTION METHODS

Sr. No.	Method	Reduction in encoding time	Bit rate increment
1	[4]	45%	0.19%
2	[5]	15%	0.13%
3	[6]	35.51%	0.45%
4	[7]	45%	-
5	[8]	41.4%	1.88%
6	[9]	25%	0.16%
7	[10]	53%	1.32%
8	[11]	43.3%	1.5%
9	[12]	50%	0.6%
10	[13]	40%	1.42%
11	[14]	-	-
12	[15]	-	-
13	[16]	41%	1.9%
14	[17]	53%	2.74%
15	[18]	46.15%	1.2%

IV. CONCLUSIONS

In this paper, the fast inter CU decision algorithms for HEVC encoding is surveyed. The methods surveyed above are compared in terms of encoding time reduction and RD performance. All the above methods provide significant reduction in encoding time but also have negligible coding loss. In the future, the algorithms surveyed in this paper can be further optimized and enhanced.

V. ACKNOWLEDGMENT

I am indeed thankful to my guide Prof. S. S. Agrawal for her able guidance and assistance to complete this paper; otherwise it would not have been accomplished. I extend my special thanks to Head of Department of Electronics & Telecommunication, Dr. S. K. Shah who extended the preparatory steps of this paper-work. I am also thankful to the head & Principle of STESS, SMT. Kashibai Navale College of Engineering, Dr. A.V. Deshpande for his valued support and faith on me.

REFERENCES

[1] G. 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. 1649–1667, Dec. 2012.

[2] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," IEEE Signal Process. Mag., vol. 15, no. 6, pp. 74–90, Nov. 1998.J.

[3] A. Lee, D. S. Jun, J. Kim, J. S. Choi, and J. Kim "An Efficient Inter Prediction Mode Decision Method for Fast Motion Estimation in High Efficiency Video Coding," *ETRI Journal*, to be published.

[4] Leng, L. Sun, T. Ikenaga, and S. Sakaida, "Content based hierarchical fast coding unit decision algorithm for HEVC," *in Proc. Int. Conf. Multimedia and Signal Processing*, vol. 1, pp. 56–59, May 2011.

[5] L. Shen, Z. Liu, X. Zhang, W. Zhao, and Z. Zhang, "An effective cu size decision method for HEVC encoders," *IEEE Trans. Multimedia*, vol. 15, no. 2, pp. 465–470, Feb. 2013.

[6] J. Kim, J. Yang, K. Won, and B. Jeon, "Early determination of mode decision for HEVC," *in Proc. Picture Coding Symp. (PCS)*, pp. 449–452, May 2012.

[7] M. Cassa, M. Naccari, and F. Pereira, "Fast rate distortion optimization for the emerging HEVC standard," *in Proc. Picture Coding Symp. (PCS)*, pp. 493–496, May 2012.

[8] X. Shen, L. Yu, and J. Chen, "Fast coding unit size selection for HEVC based on bayesian decision rule," *in Proc. Picture Coding Symp. (PCS)*, pp. 453–456, May 2012.

[9] Y. Zhang, H. Wang, and Z. Li, "Fast coding unit depth decision algorithm for interframe coding in HEVC," in Proc. Data Compression Conf., pp. 53–62, 2013.

[10] J. Xiong, H. Li, F. Meng, B. Zeng, S. Zhu, and Q. Wu, "Fast and efficient inter CU decision for high efficiency video coding," in Proc. IEEE Int. Conf. Image Process., to be published.

[11] X. Shen and L. Yu, "CU splitting early termination based on weighted SVM," EURASIP J. Image Video Process., vol. 2013, no. 1, pp. 1–11, 2013.

[12] S. Cho and M. Kim, "Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding," IEEE Trans. Circuits Syst. Video Technol., vol. 23, no. 9, pp. 1555–1564, Sep. 2013.

[13] G. Correa, P. Assuncao, L. Agostini, and L. A. da Silva Cruz, "Complexity control of HEVC through quadtree depth estimation," in Proc. IEEE EUROCON, pp. 81–86, Jul. 2013.

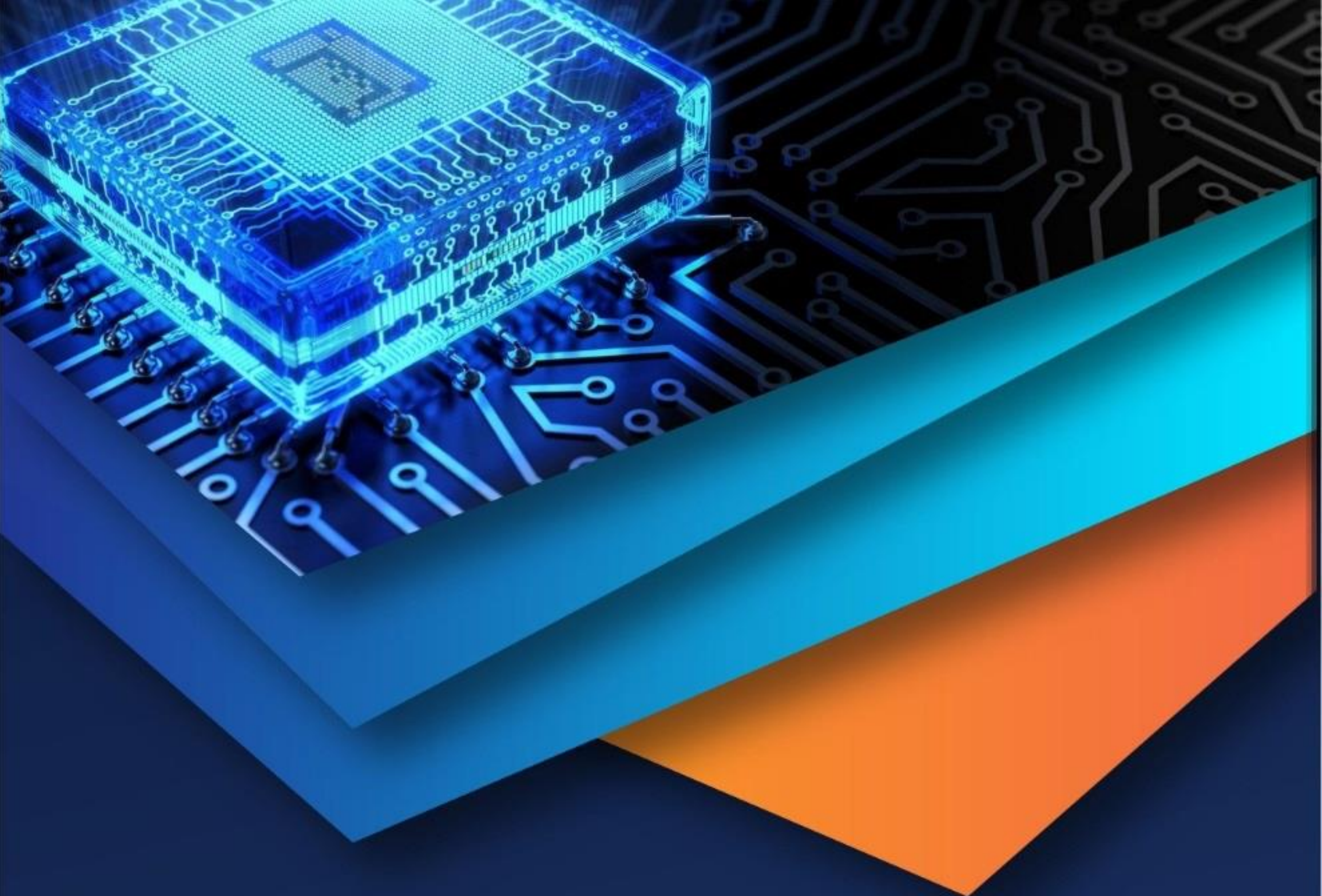
[14] T. Zhao, Z. Wang, and S. Kwong, "Flexible mode selection and complexity allocation in high efficiency video coding," IEEE J. Sel. Topics Signal Process., vol. 7, no. 6, pp. 1135–1144, Dec. 2013.

[15] J. Kim, S. Jeong, S. Cho, and J. Choi, "Adaptive coding unit early termination algorithm for HEVC," in Proc. IEEE ICCE, pp. 261–262, Jan. 2012.

[16] J. Xiong, H. Li, Q. Wu, and F. Meng, "A Fast HEVC Inter CU Selection Method Based on Pyramid Motion Divergence," IEEE trans. on multimedia, vol.16, no.2, pp.559 - 564, Feb.2014.

[17] J. Xiong, H. Li, S. Zhu, Q. Wu, and B. Zeng, "MRF-Based Fast HEVC Inter CU Decision With the Variance of Absolute Differences," *IEEE trans. on multimedia,* vol.16, no.8, pp.2141 - 2153, Dec.2014.

[18] S. Ahn, B. Lee and M. Kim, "A Novel Fast CU Encoding Scheme based on Spatio-Temporal Encoding Parameters for HEVC Inter Coding," *IEEE trans. on Circuits and Systems for Video Technology*, to be published.



10.22214/IJRASET

INDEX COPERNICUS

45.98

ISRA JIF

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY