

Bit Rate Transcoding for High Efficiency Video Coding



Sonal K. Jagtap, Kanchan L. Dombale, Himali B. Ghorpade

Abstract: High efficiency video coding (HEVC) has demonstrated a notable increase in compression performance and is taken as a successor to H.264/AVC. Efficient bit rate adaptation algorithms are required to contain the HEVC standard between real life community facilities. A present issue of bit rate transcoding is its high computational complexity which is related with the encoder of a cascaded pixel domain transcoder. This paper gives Top to Bottom (T2B) approach to reduce complexity by using different complexity schemes. Proposed approach is effective in reducing complexity in Coding Unit (CU) optimization level. Coding Unit has been analyzed in T2B Approach. While examining the coding unit information of the input video is turned to account for decreasing the number of evaluation and early terminate the process. For the Prediction Unit (PU) level the units are powerfully chosen contingent upon likelihood of Prediction Unit sizes and co-found input prediction partitioning. By utilizing this approach, complexity scalable bit rate transcoding has achieved. Machine learning approach can be used to control computational complexity. Additionally, the T2B strategy is able to gain a spread on trade-offs in transrating complexity and coding performance. Using T2B approach 15% encoding time saving is accomplished. From this scheme, for the less resolution video 27% time saving has achieved.

Keywords: Coding Unit (CU), High Efficiency Video Coding (HEVC), Transcoding, Video Coding.

I. INTRODUCTION

HEVC is the latest standard which is developed by the Video Coding Expert Group (VCEG) and the Moving Picture Expert Group (MPEG). HEVC saves around 50% bit rate as compared to its previous standard i.e. H.264. Because of high compression efficiency HEVC is widely utilized in many applications. To use HEVC in daily life applications some transcoding algorithms are required. Due to transcoding network bandwidth constraints can be adjusted and parameters of the transmitter side can be changed as per receiver requirements [1]. Transcoding are of two types 1) Homogeneous Transcoding 2) Heterogeneous Transcoding .In homogeneous video transcoding the transmitter bitstream

parameters are changed according to the receiver bitstream parameters.

Open and closed loop transcoding are the two approaches of transcoding. Open-loop transcoding is non recursive type of transcoding and closed loop transcoding is recursive type of transcoding. Open loop transcoder are less complex as compared to closed loop transcoders.

The coefficients are transformed in to the frequency domain and are further quantized in open loop transcoding, and in open loop transcoding motion vectors are not evaluated.

The difference between encoder and decoder causes drift. This drift can be reduced by employing closed loop cascaded pixel domain transcoder. In this transcoder bitstream is initially decoded and then further encoded to match the receiver bit rate. Because of large re-encoding evaluations, a high complexity arises. To convert information from one type to another encoder is being used. It transforms analog signal into digital signal. To change information from digital to analog decoders are used. Encoded signal can be recovered back by using decoder. In another option of requantization dynamic rate shaping removes high frequency coefficients. Different bit rate changing techniques have been given for H.264 bit streams.

A successful joined transrating architecture containing a less complexity mixed with a drift dropping closed loop technique. A bit rate transcoding method uses requantization of the transform coefficients integrated in a rate control algorithm. The handiness of an open-circle transcoder is its less intricacy, however the drift influences yields in visual quality misfortunes as a result of the error propagation. This drift can be downsize in a closed-loop cascaded pixel domain transcoder where the video bitstream is decoded and re encoded to coordinate the objective bit rate [3].

In this paper the T2B approach is given to reduce transrating complexity of HEVC. The rate distortion (RD) cost is used to reduce encoding time. The maximum size of a Coding Tree Unit (CTU) is 64×64 pixels and is further splitted into Coding Unit (CU). For CU's skip, inter and intra prediction modes are supported. Splitting is done process for CUs from depth 0 which is 64×64 pixel CU to depth 3 which is 8×8 pixel CU. Every coding unit is the source for farther evaluation of the prediction unit (PU), and transform unit (TU).

The further part of this paper is organized as follows. In Section II, related work is introduced. The bit rate transcoding architecture, methodology and flowchart are discussed in Section III. Section IV presents simulation results. Section V discusses the conclusions and future work.

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Dr. Sonal. K. Jagtap*, Doctorate (Ph.D) Degree, Electronics Engineering, Shivaji University, Kolhapur (Maharashtra) India.

Kanchan Laxman Dombale, B.E Degree, Electronics and Telecommunication Engineering, Dhole Patil College of Engineering, Pune (Maharashtra) India.

Himali Babanrao Ghorpade, B.E Degree, Electronics and Telecommunication Engineering, Sanjay Ghodawat Group of Institutes, Kolhapur (Maharashtra) India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

II. RELATED WORK

In the first section homogeneous video transcoding is explained. In homogeneous video transcoding the format of video coding does not get changed. And in the second section Heterogeneous video encoder optimization strategies are described. There have been a couple of efforts performed for erection the higher video coding. The homogeneous video transcoding methods and research difficulties explained between [1],[8] which gives the quick concept regarding transcoding techniques in same video coding format in which we can change bit rate, spatial resolution, temporal resolution etc.

For fast transcoding machine learning approach is used, which gives effective results in transrating [2], [3]. Bit rate transcoding approach for downsizing the pre-encoded video using the Machine learning based principle is given in [4]. Motion estimation approach for compression of H.264 explained in [5], In which motion in video is used for compression. By skipping particular frames the compression can be achieved. The ways for skipping frames are in [6]. Using RD optimization frame skipping is in [7], which is basically implemented for low bit rate. Fast transcoding for video transmission is explained in [8].

In heterogeneous transcoding format of the video coding standard gets changed. For heterogeneous transcoding there are different algorithms. Many algorithms has been designed in conformity with expand the speed of the HEVC encoder. These algorithms turn to advantage the field yet use the temporal/spatial association within the video in conformity with recommend the coding information over a coding unit. A quickly coding unit depth selection approach and different modes like inter and intra approaches used are explained in [9]. Data mining is also used in HEVC encoding and the approach is given in [10]. A speedy coding unit size selection approach is described within [11]. Another direction may result greater complexity reduction because it utilizes the relation between the coding information between the input and output video. Heterogeneous transcoding based on dynamic thresholding and content modeling is given, where we utilize the co-relation in between input and output video information, and by using information we convert H.264 to HEVC [12].

III. TRANSCODING ARCHITECTURE AND METHODOLOGY

The transcoding architecture consists mainly two stages decoder and re-encoder. CU splitting decision block is the main working area in this method. Input to the HEVC transcoder is the HEVC video .Input HEVC video is then decoded for the video information. Decoded information can be used for the re-encoding in the transcoding stage. A decoder is an algorithm that converts information from one format to another format. The HEVC decoder takes the input as an encoded HEVC bitstream and decodes it into pixel format. Reconstructed frames are divided into coding tree unit (CTU) and CTU's are divided into CU's.

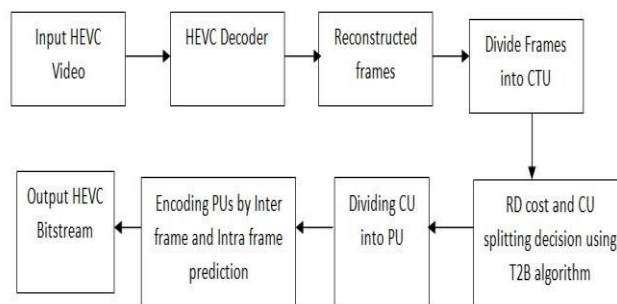


Fig.1. Transcoding architecture

Top 2 bottom (T2B) approach is used for CU splitting decision. For taking decision whether CU should be split or not the rate distortion (RD) cost is found. First for maximum depth the RD cost found (RD Not split), then split CU with current depth level. Find RD cost of all splitted CU's (RD split) and sum up them if the RD not split cost is less than RD split then discard the further splitting. If RD not split cost is greater than RD split then split the CU with next depth. CU's are then divided into PU's. PU's are encoded by inter frame and intra frame prediction.

A. Finding RD Cost Function

The RD cost of each CTU and CU is calculated to take decision of CTU and CU splitting. The sum of square difference is calculated between original block and reconstructed block using following equation,

$$D = [\text{abs}(\text{Original} - \text{Reconstructed})]^2 \quad (1)$$

Where, D is the distortion, Original and Reconstructed are original and reconstructed CU. RD cost is calculated by following formula

$$\text{RD cost} = \text{Distortion} \times \text{Bitrate} \quad (2)$$

For variable bitrate need to find the cost function using the Lagrangian multiplier. Now estimation of Lagrangian cost is done by using function as;

$$C = D + \lambda_{\text{mode}} \cdot B \quad (3)$$

Where,

$$\lambda_{\text{mode}} = 0.85 * 2^{(QP-12)/3} \quad (4)$$

B= no. of bits, QP is quantization parameter and λ_{mode} is Lagrangian multiplier.

Now each block is divided in to four equal size blocks of 32 x 32 and then SSD and rate distortion cost of each sub-block is calculated. If the addition of rate distortion cost of four 32 x 32 block is less than the rate distortion cost of original 64 x 64 block then splitting of block is confirmed and split mode is decided for that block. Same procedure is carried out for depth 1 (32x32) and depth 2 (16x16).

B. FLOWCHART

The T2B approach is explained through flowchart in the fig. below. It starts with selecting maximum depth. After selecting the maximum depth the RD cost of the current CU depth is calculated then the CU is partitioned is done till 4*4 blocks. The RD cost of each sub block is calculated and the summation is taken of these RD costs. If RD cost of original is minimum then splitting is discarded and if the RD cost is more that time splitting is done. Then check for the lower depth CU.



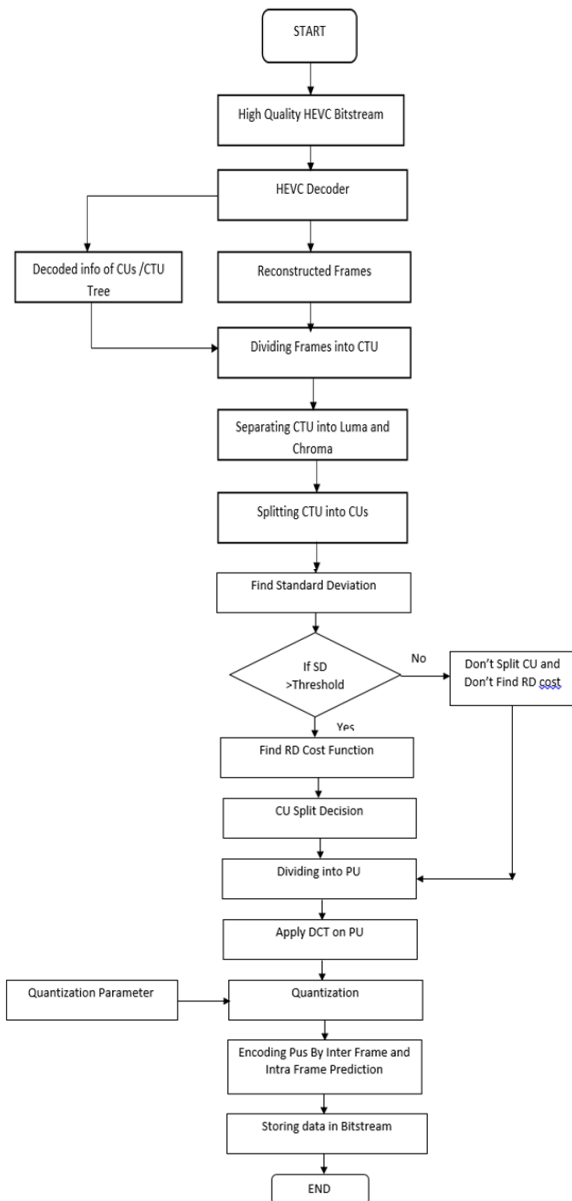


Fig.2.Flowchart of RD cost

The minimum depth for CU is 64*64 and maximum depth for CU is 4*4. To encode maximum depth CU required encoding time greater than the encoding time required to encode the minimum depth.

IV. RESULTS AND DISCUSSION

To evaluate the proposed method the modified algorithm is applied on six different video sequences of different resolutions. The results can be calculated for the constant bitrate as well as variable bit rate. In this simulation variable bitrates are used. For different quantization parameters results are evaluated. The simulation is tested on ten frames of each video sequence. Results are calculated to provide answers to the following questions .How much time required for the frame splitting? How much time required for the encoding the frame?

The simulation is performed on 10 frames of each video sequence, which has different resolution and bit rate. To get answer of the first question the three algorithms are applied on the different video sequences. First algorithm is without split decision approach, second is Top 2 Bottom approach.

Table.1. Specifications of standard video database

Video Sequence	Resolution(HXW)	Frame Rate (fps)
1. Akiyo	128 X 192	30
2. BQ Square	240 X 416	30
3. BasketBall Pass	288 X 352	30
4. BQ Mall	480 X 832	30
5. Four People	704 X 1280	30
6. Tennis	1056 X 1920	30

Simulation is done on six different videos of different resolution as, Akiyo, BQ Square, BasketBall Pass, BQ Mall, Four people and Tennis. BasketBall Pass sequence is chose for the further simulation which is of 352 X 288 resolution.



Fig.5.Different video sequences used for experimentation [3]

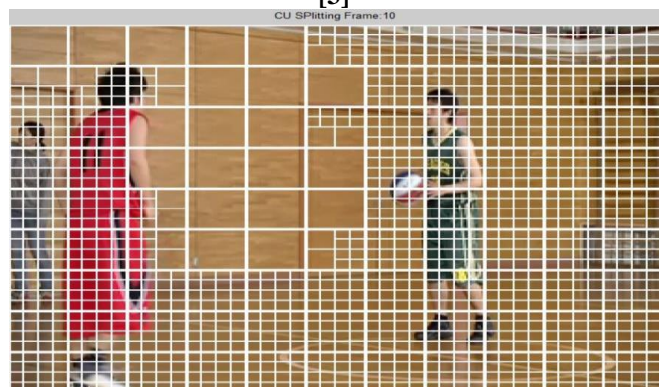


Fig.6.Splitted basket ball video sequence using T2B approach

CU splitting is done where maximum motion and texture difference is present. Where less or no motion present and there is no texture difference no need to split the CTU and CU. From the frame encoding time the time savings has found. In the T(without split decision) without applying any algorithm the encoding time has calculated and in T(T2B) by applying T2B algorithm Encoding time has calculated.

By using below formula the total time savings in % has calculated.

$$\text{Time saving(\%)} = \frac{T(\text{Without split decision}) - T(\text{T2B})}{T(\text{Without split decision})} \quad (5)$$

The average encoding time required using without split decision approach is 176.72 Sec and average encoding time using Top 2 Bottom algorithm is 150.12 Sec. For the less resolution videos maximum encoding time savings has achieved.

Table.2. Comparison results for T2B algorithm

Algorithm	Without split decision	T2B split decision algorithm	Encoding time saving (%)
Video Sequence	Encoding Time(Sec)	Encoding Time(Sec)	Encoding Time saving (%)
Akiyo	9.5581	6.9928	26.83
BQ Square	28.4447	23.6963	16.69
Basketball Pass	30.5523	26.7705	12.37
BQ Mall	117.8627	103.8816	11.86
Four People	270.7316	226.520	16.33
Tennis	603.1827	512.8592	14.97
Average	176.72	150.12	15.05

The encoding time for Akiyo video sequence without split decision approach is 9.5581 sec and by using T2B approach it is 6.9928 for 10 frames. The throughout experiment is performed on 10 frames of each video sequence.

V. CONCLUSION

T2B approach gives around 15 % time saving as compared to without split decision approach. In less resolution and bit rate videos we get maximum time saving. As shown in Table 1 in Akiyo video sequence maximum time saving that is 27% achieved. And in high resolution videos around 15 % time savings achieved. By comparing T2B approach with without split decision approach around 15% Encoding time savings has been achieved. For higher resolution videos the encoding time saving is less as compared to less resolution videos. Interconnection between co-located PUs and motion estimation information can be utilized to further decrease the complexity of transrating as well for more encoding time reduction machine learning algorithms can be used.

REFERENCES

1. G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, pp. 1649–1668, May 2012.
2. L.P. Van, J. De Cock, G. Van Wallendael, S. Van Leuven, R. RodriguezSanchez, J. Martinez, P. Lambert, and R. Van de Walle, "Fast transrating for high efficiency video coding based on machine learning," *Proc. IEEE Int. Conferenc Image Processing*, pp. 1573–1577, Sep. 2013.
3. Luong Pham Van, Johan De Praeter, Glenn Van Wallendael, Sebastiaan Van Leuven, Jan De Cock "Efficient Bit Rate Transcoding for High Efficiency Video coding" *IEEE transactions on multimedia*, Vol.18, No.3, March 2016
4. Luong Pham Van; Johan De Praeter; Glenn Van Wallendael; Jan DeCock RikVande, Walle, "Machine Learning for arbitrary downsizing of pre-encoded video in HEVC" *IEEE International Conference on Consumer Electronics (ICCE)*, Pages: 406 – 407, Year: 2015.
5. Ying Zhang; Jiao Wang; Sorina Dumitrescu, "Efficient motion estimation scheme for downsizing H.264 Homogeneous video" *,IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Pages: 520 - 523, Year: 2011.

6. Siyoung Yang; Donghyung Kim; Yeonggyun Jeon; Jechang Jeong, "An efficient motion re-estimation algorithm for frame-skipping video transcoding" *,IEEE International Conference on Image Processing*, vol: 3 ,Pages: III - 668-71, , Year: 2005.
7. Goo-rak Kwon; Sang-hee Park; Jae-won Kim; Sung-jea Ko, "Real time RD optimized Frame skipping Transcoder for low bit rate video transmission" *The Sixth IEEE International Conference on Computer and Information Technology (CIT'06)*, Pages: 177 - 177, Year: 2006.
8. G. Van Wallendael, J. De Cock, and R. Van de Walle, "Fast transcoding for video delivery by means of a control stream," *Proc. IEEE Int. Conf. Image Process.*, pp. 733–736, Sep. 2012
9. L. Shen, Z. Zhang, and P. An, "Fast CU size decision and mode decision algorithm for HEVC intra coding," *IEEE Trans. Consum. Electron.* vol. 59, no. 1, pp. 207–213, Feb. 2013.
10. G. Correa, P. Assuncao, L. Volcan Agostini, and L. da Silva Cruz, "Fast HEVC encoding decisions using data mining," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 4, pp. 660–673, Apr. 2015.
11. J. Lee, S. Kim, K. Lim, and S. Lee, "A fast CU size decision algorithm for HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 3, pp. 411–421, Mar. 2015.
12. E. Peixoto, T. Shanableh, and E. Izquierdo, "H.264/AVC to HEVC video transcoder based on dynamic thresholding and content modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 1, pp. 99112, Jan. 2014.

AUTHORS PROFILE



Dr. Sonal. K. Jagtap received her Doctorate (Ph.D) degree in Electronics Engineering from Shivaji University, Kolhapur in 2014. She received M.E. degree with specialization in Microwave Engineering from COEP, Pune University in 2002 and B.E. degree in Electronics and Telecommunication Engineering in 1999. Her research interest includes video and image processing, Artificial Intelligence and Machine Learning. She is currently working as Professor and Head, Dept of E&TC Engineering in Smt. Kashibai Navale College of Engineering, Pune-41



Kanchan Laxman Dombale received B.E degree in Electronics and Telecommunication Engineering from Dhole Patil College of Engineering, Pune in 2014. She has completed M.E .degree with specialization in Signal Processing in 2017 from Smt. Kashibai Navale College of Engineering, Pune.



Himali Babanrao Ghorpade received B.E degree in Electronics and Telecommunication Engineering from Sanjay Ghodawat Group of Institutes ,Kolhapur in 2015. She has completed M.E .degree with specialization in Signal Processing in 2017 from Smt. Kashibai Navale College of Engineering , Pune.

