International Journal of Modern Trends in Engineering and Research
www.ijmter.com
e-ISSN No.:2349-9745, Date: 2-4 July, 2015

Heterogeneous Video Transcoder for H.264/AVC to HEVC: Review

S. J. Kaware[1],S. K. Jagtap[2]

[1]*Department of E&TC, Smt. Kashibai Navale College of Engineering Vadgaon (Bk), Pune-41, India, shri.kaware@gmail.com*
[2]*Departmentof E&TC,Smt. Kashibai Navale College of Engineering Vadgaon (Bk), Pune-41, India, skjagtap.skncoe@sinhgad.edu*

Abstract— As a successor to the H.264/Advance Video Coding (AVC), High Efficiency Video Coding **(**HEVC) is invented to get more compression and high quality video. But HEVC increases the complexity of video coding. To reduce this complexity various methods are proposed with different approach which also enhance the video compression and quality of video. This paper contains different techniques for video transcoding such as dynamic thresholding, mode mapping, machine learning, complexity scalable and background modeling. Comparison of all techniques gives their advantages and disadvantages to modify video transcoder which benefits from the strength points of existing methods.

Keywords- background modeling, complexity scalability, content modeling and early termination, dynamic thresholding, HEVC, machine learning, mode mapping.

I. INTRODUCTION

Different areas for digital video technology which was not previously covered by other video coding standard, has been enabled by H.264/AVC standard. The H.264/AVC is very successful and it is widely used for many applications like for broadcasting HD TV signals over satellite, internet and mobile network video, real time conversation application i.e. video chat, video conferencing etc. However, due to the increase in both diversity of services and popularity of High Definition (HD) video, the need of more efficient video coding standard superior to H.264 capabilities become stronger. Due to the amendment of higher resolution in stereo or multiview devices the need become even stronger. But, there are several needs for video transcoding. To meet the specific network requirements such as bitrate, resolution, frame rate etc., a previously encoded video data may need to be transcoded to new standard.
HEVC was introduced by theJoint Collaborative Team on Video Coding(JCTVC) group as a new video coding standard to achieve better compression rate than H.264/AVC. HEVC focus on two main issues i.e. increased video resolution and increased use of parallel processing architectures. HEVC has generic syntax and suited for all applications of H.264/AVC. HEVC has standardized bit stream structure and syntax. To learn how to use the standard, the standardization effort not only includes the development of a text specification document, but also reference software source code as an example of how HEVC video can be encoded and decoded. The draft reference software has been used as a research tool for the internal work of the committee during the design of the standard, and can also be used as a general research tool and as the basis of products [1].The two main categories of transcoders are homogeneous transcoding (the conversion of bit streams within the

same format) and heterogeneous transcoding (i.e., between different formats). The transcoding from H.264/AVC to HEVC falls under heterogeneous transcoding category [2].
The flow for the paper is described as below: Section II, related work of various authors is mention. Sections III, different techniques used for transcoding from H.264 to HEVC are described. Section IV, comparison of transcoding techniques for H.264/AVC to HEVC is given. In section V, conclusions are discussed. Future work is proposed in section VI.

II. RELATED WORK

Various authors [2-7] published various results of different transcoding techniques for HEVC. Transcoding technique based on dynamic thresholding [2] is described to increase the speed of transcoding. Transcoding based on mode mapping [3] is implemented to get more stable results over wide range of sequences. As there are very large numbers of modes in HEVC, mode mapping is important technique of transcoding. Machine learning technique is used [4] to achieve higher speed for transcoding. For complexity reduction, the new codec based on complexity scalable technique is introduced [5]. Reduction in transcoding time is done by using coding unit classification technique with background modeling [6]. A fast HEVC transcoder based on content modeling and early termination is explained in [7].

III. TECHNIQUES FOR TRANSCODING

Different approaches for transcoding of H.264/AVC to H.265/HEVC are shown in fig. 1. It focuses on dynamic thresholding, mode mapping, machine learning, complexity scalable and background modeling, content modeling and early termination.

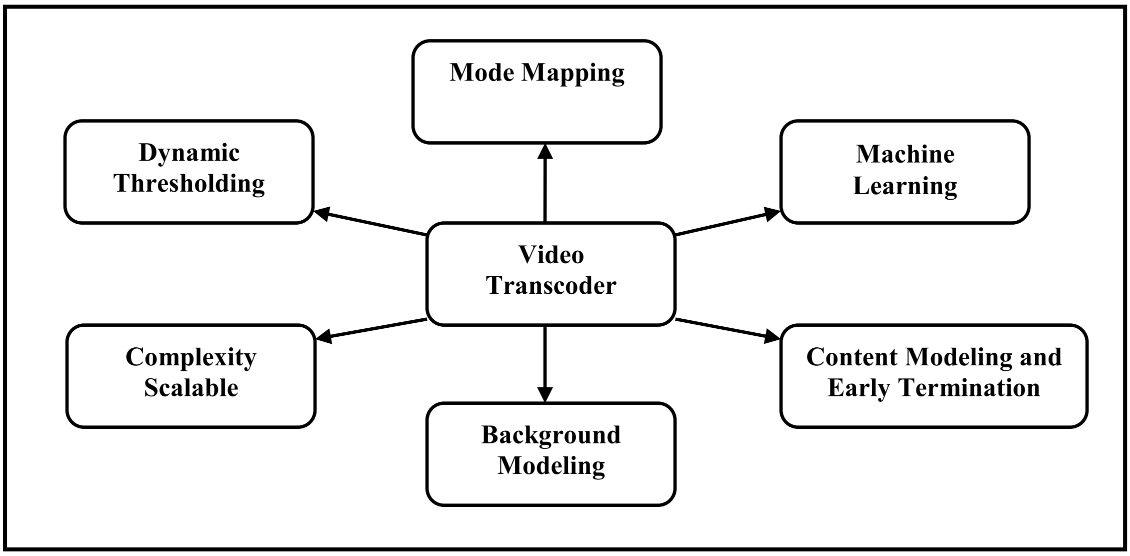

Figure 1. Different approaches for transcoding

Transcoding technique based on dynamic thresholding is described in [2]. In this method, transcoding is done by extracting different features like motion vector variance distance, motion vector variance phase, number of DCT coefficient, and energy of DCT coefficient of incoming

H.264/AVC. Two threshold i.e. T_{low} and T_{high} are used to decided mode of CU and PU.There are two stages of video coding i.e. training stage and transcoding stage. Training stage is applied only on K frames and all modes of CU and PU are tasted. Whereas transcodingstage is applied on N-K frames, where N is the total number of frames of video, and modes are mapped from training stage. This approach increases the transcoding speed at the cost of rise in complexity.

Transcoder for H.264/AVC to HEVC based on mode mapping technique is explained in [3]. In this method, mapping of modes of H.264/AVC macroblocks into HEVC coding unit (CU's) is done by using machine learning technique. The HEVC codec uses a very large number of modes, making mode mapping algorithm the most important part for this transcoder. To build a machine learning model, two approaches are evaluated. First is the static training, in which the model is created offline and this model is used for transcoding any video sequence. Second approach is dynamic training, in which the model is created at run time. In this approach two stages are used i.e. training stage and transcoding stage. In training stage, information in the incoming bit stream of H.264/AVC is gathered by performing full re-encoding. This collected information is used to map the modes for HEVC. The main advantage of this method is that using the dynamic training yields more stable results over a wide range of sequences but there is small increase in complexity.

Fast H.264/AVC to HEVC transcoding based on machine learning is proposed in [4]. The trivial solution to transcoding problems is to fully decode the source bit stream and completely re-encode it in the target codec. This procedure is usually referred as the trivial transcoder. Rate distortion (RD) performance is used to achieve the best results. But it is a time consuming task. To speed up the transcoding technique, the previously encoded information, such as the prediction modes, motion information and encoded residuals, can be used to re-encode the video data. This is even more critical in high bitrate video, due to the large amount of data involved.So proposed a transcoder uses machine learning techniques in order to decide which HEVC modes will be tested. It uses features extracted from the H.264/AVC bit stream, such as the motion vectors and DCT coefficients, to build the machine learning model. The main contributions of this are a more robust feature selection for the training process, and a better CU classification strategy that targets faster transcoding. This method has advantage of higher speed up to 1.65 times faster than other transcoder but it has a higher loss in bitrate.

H.264/AVC to HEVC codec based on complexity scalable is focused in [5]. The transcoding techniques proposed here aims primarily to reduce the complexity of the MV Reuse transcoder. It achieves this mainly by testing less CUs and PUs than trivial transcoder. The proposed transcoder uses a similarity metric, the MV Variance Distance and, according to this metric, it makes the decision of which PUs will be tested for a given CU and whether this CU will be split or not. The similarity metric produces a value $\upsilon \geq 0$ for each CU in the HEVC, and two thresholds T_{low} and T_{high} are used to decide how that CU will be tested. Starting from the largest size CU, the transcoder divides the possible partitions units (PUs) to be tested in 4 groups: (i) SKIP; (ii) inter 2N×2N; (iii) all remaining inter modes (2N×N, N ×2N, the AMP modes, and N×N); and (iv) the intra modes (2N×2N, N×Nand PCM). In addition, the transcoder can decide if the CU will be split or not (if so, the CU is split in four sub-CUs, as usual). When a CU is split, the algorithm is applied in the same manner to each of the four sub-CUs, computing a new similarity value υ for the sub-CUs, until the final possible depth is reached. Note that the SKIP mode is always tested, even if the CU is considered dissimilar, as the complexity to test the SKIP is small, compared to the other modes. Furthermore, the intra modes are only tested if some part of this CU was coded as intra in the H.264/AVC bit stream, and the default fast mode decision algorithms in the HEVC still applies (for testing the intra PCM and to decide for early skip).Complexity reduction is the main advantage of this method but it has increase in rate distortion.

Transcoding of AVC to HEVC based on coding unit classification with background modelling is described in [6].This technique is used to obtain a high efficiency and low complexity surveillance video. It uses background modeling to get high efficiency, where the modeling of background frame is done by using originally decoded frames. Then, for enhancement of background prediction efficiency, the transcoding of modeled background frame is done to get into HEVC stream. To get low complexity, firstly algorithm for a CU classification using the decoded motion vectors is developed and the modeled background frames are given as input. As an outcome, three types of the decoded data are generated. First is the background CUs (BCs), which are mainly background pixels, second is foreground CUs (FCs), which are mainly foreground pixels and third is hybrid foreground and background CUs (HCs). Every CU category has different statistics. Depending on which each transcoded CU has its own different characteristics. In the recursive coding structure of HEVC, for CU partitioning, there are different terminated depths and for inter prediction, there are different PU patterns. This method reduces transcoding time but there is not sufficient utilization of decoding information.

A fast HEVC transcoder based on Content Modeling and Early Termination is explained in [7]. This method is based on two main modules: a coding unit (CU) classification module that relies on a machine learning technique in order to map H.264/AVC macroblocks into HEVC CUs; and an early termination technique that is based on statistical modeling of the HEVC rate-distortion (RD) cost in order to further speed-up the transcoding. The transcoder is built around an established two stage transcoding. In the first stage, called the training stage, full re-encoding is performed while the H.264/AVC and the HEVC information are gathered. This information is then used to build both the CU classification model and the early termination sieves that are used in the second stage (called the transcoding stage). The proposed method is 3.83 times faster, on average, than the trivial transcoder, and 1.8 times faster than a previous transcoding solution, while yielding a RD loss of 4% compared to this solution.

IV. COMPARISON AND PERFORMANCE STUDIES

Different transcoding techniques have variations in transcoding speed to perform video coding. Table 1 showscomparison of transcoding techniques for H.264/AVC to HEVC related to more or less loss in bitrate. Higher speed is achieved at the cost of bitrate loss.

Table 1. Comparison of transcoding techniques for H.264/AVC to HEVC

Techniques For Transcoding	Dynamic Thresholding	Mode Mapping	Machine Learning	Complexity Scalable	Background Modeling	Content Modeling And Early Termination
Speed	2.15	2.26	3.4	4.13	-	1.8
Loss in bitrate (%)	5.42	3.6	8.77	8.24	3.7	4

Mode mapping method is 2.26 times faster than trivial transcoding but it has 3.6% loss in bit rate whereas machine learning technique is much faster than previous method but increase in bitrate loss. Method like complexity scalable focuses on transcoding speed which is faster than others but loss in bitrate also increased.

V. CONCLUSIONS

Transcoding speed differs with technique chosen to transcode the video sequence. Increase in speed results in increase in loss of bitrate. So modified video transcoder can be proposed which benefits from the strength points of existing methods and accordingly improve the rate distortion performance of H.264 which allow for new applications, such as beyond high definition resolutions.

VI. FUTURE SCOPE

For more perceptual video quality, compression performance can be improved comparative to existing standards by reducing the rate distortion loss. Different set of features could be explored to improve the transcoder performance. Other machine learning techniques could be used.

ACKNOWLEDGMENT

I am indeed thankful to my guide Prof.(Dr.) S. K. Jagtap for her able guidance and assistance to complete this paper. Author and co-author of this paper also would like to thank other researchers who provided the platform for this paper.

REFERENCES

[1] G. J. Sullivan, J. Ohm, H. W. Jin, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," IEEE Trans. Circuits Syst. Video Technology, vol. 22, no. 12, pp. 1649–1668, Jul. 2012.
[2] E. Peixoto, T. Shanableh, and E. Izquierdo, "H.264/AVC to HEVC video transcoder based on dynamic thresholding and content modeling," IEEE transactions on Circuits And Systems for Video Technology, Vol. 24, No. 1, Jan. 2014.
[3] E. Peixoto, B. Macchiavello, E. M. Hung, A. Zaghetto, T. Shanableh, E. Izquierdo, "An H.264/AVC to HEVC video transcoder based on mode mapping," in Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 1972–1976, 2013.
[4] E. Peixoto, B. Macchiavello, E. M. Hung, "Fast H.264/AVC to HEVC transcoding based on machine learning," in Proc. International Telecommunications Symposium (ITS), pp. 1-4, 2014.
[5] E. Peixoto and E. Izquierdo, "A complexity scalable transcoder from H.264/AVC to the new HEVC codec," in Proc. IEEE Int. Conf. Image Process. (ICIP), pp. 737–740, Sep. 2012.
[6] P. Xinga, Y. Tian, X. Zhang, Y. Wang, T. Huang, "A coding unit classification based AVC to HEVC transcoding with background modeling for surveillance videos," in Proc. IEEE Int. Conf. Visual Communications and Image Processing (VCIP), pp. 1-6, 2013.
[7] E. Peixoto, B. Macchiavello, E.M. Hung , R.L. de Queiroz,"A fast HEVC transcoder based on content modeling and early termination," IEEE Int. Conf. Image Process. (ICIP), pp. 2532–2536, 2014.

