International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 3 Issue: 6 4254 – 4256

Heterogeneous Video Transcoder for H.264/AVC to HEVC

S. J. Kaware
Department of E&TC
Smt. Kashibai Navale College of Engineering
Vadgaon (Bk), Pune-41, India
shri.kaware@gmail.com

S. K. Jagtap
Department of E&TC
Smt. Kashibai Navale College of Engineering
Vadgaon (Bk), Pune-41, India
skjagtap.skncoe@sinhgad.edu

Abstract— The new video coding standard, High Efficiency Video Coding, was developed to succeed the current standard, H.264/Advance Video Coding. However, there is a lot of legacy content encoded with H.264. So the new efficient method is proposed for transcoding the H.264 encoded video into high efficiency video coding format. In proposed method, two stages are implemented. In training stage, transcoding is done using SSD method and different coding parameters or features are extracted from incoming H.264. In transcoding stage, the best mode of outgoing coding unit partitions are decided by calculating threshold value and optimum weight using extracted features. Then it is evaluated by doing experiments on different videos.

Keywords- *Coding unit (CU), Energy of DCT coefficient, High Efficiency Video Coding (HEVC), Motion vector variance, Optimum weight vector, Thresholding, Video transcoding.*

*****

I. INTRODUCTION

Now days, there is increasing need of high quality videos due to popularity of HD and UHD videos. But there are several challenges on today's network. So to overcome these issues, the new video coding standard is proposed to enhance the current standard, H.264/AVC.

H.265/HEVC (High Efficiency Video Coding) was designed by the JCTVC group to replace the current H.264/AVC standard. The improvement of the rate distortion performance of H.264/AVC is the main intention of the HEVC codec. HEVC codec provides about twice improvement in compression compared to AVC. This will enable the video cloud servers (like YouTube) to store more video data over less memory and also provides ease for database management. Transcoding is the process of converting one compressed bitstream to another compressed bitstream.

The flow for the paper is described as below: Section II, literature survey is described by related work of various authors and comparing different techniques used for transcoding from H.264 to HEVC. Section III, proposed method for transcoding for H.264/AVC to HEVC is given. In section IV, experimental results are given. Section V, conclusions are discussed.

II. LITERATURE SURVEY

HEVC is the latest topic for researcher and lot of work is going on for the improvement of coding standard. Different researchers proposed different approaches to transcode the video in HEVC format. Some related work is given below.

A. Related Work

In [2], H.264/AVC to HEVC video transcoder based on dynamic thresholding and content modeling, proposes and evaluates several transcoding algorithms from the H.264/AVC to the HEVC format. In particular, a novel transcoding architecture, in which the first frames of the sequence are used to compute the parameters so that the transcoder can learn the mapping for that particular sequence, is proposed.

In [3], an H.264/AVC to HEVC video transcoder based on mode mapping, presents a transcoding solution that uses machine learning techniques in order to map H.264/AVC macroblocks into HEVC coding units (CUs). Two alternatives to build the machine learning model are evaluated. The first uses a static training, where the model is built offline and used to transcode any video sequence. The other uses a dynamic training, with two well-defined stages: a training stage and a transcoding stage.

In [4], Fast H.264/AVC to HEVC transcoding based on machine learning, 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 a CU classification model that is used in the second stage (called the transcoding stage).

In [5], a complexity-scalable transcoder from H.264/AVC to the new HEVC codec, studies the performance of one of the most common techniques for heterogeneous transcoding, motion vector (MV) reuse, in a H.264/AVC to HEVC transcoder. The proposed transcoder is based on a new metric to compute the similarity of the H.264/AVCMVs, which is used to decide which HEVC partitions are tested on the transcoder.

In [6], a coding unit classification based AVC to HEVC transcoding with background modeling for surveillance videos; the background frame modeled from originally decoded frames is firstly transcoded into HEVC stream as long term reference to enhance the prediction efficiency.

In [7], a fast HEVC transcoder based on content modeling and early termination, 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.

ISSN: 2321-8169
4254 – 4256

B. *Comparison and Performance Studies*

Different transcoding techniques have variations in transcoding speed to perform video coding. Table 2.1 shows comparison 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 I. COMPARISON OF TRANSCODING TECHNIQUES FOR H.264/ AVC TO HEVC

Techniques For Transcoding	Speed	Loss in bitrate (%)
Dynamic Thresholding	2.15	5.14
Mode Mapping	2.26	3.6
Machine Learning	3.4	8.77
Complexity Scalable	4.13	8.24
Background Modeling	-	3.7
Content Modeling And Early Termination	1.8	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.

III. PROPOSED METHOD FOR TRANSCODING

In proposed method, for a sequence of n frames, the first k frames are used for training, and the transcoding operates in the following n–k frames. When transcoding the training sub-sequence (i.e., the first k frames), all modes in the HEVC are tested, and the H.264/AVC information is used only for training purposes. Using the information gathered at this stage, the transcoder computes the thresholds and optimal weights which are used to decide mode of HEVC for n-k frames. The advantage of using the training stage is that the transcoding parameters can be adapted to the content of the current sequence being transcoded. If the number of frames used for training is kept small, the impact on the transcoding complexity will be small as well, as the ratio (n−k/n) will be close to 1. Block diagram of proposed method is as shown in fig.1.

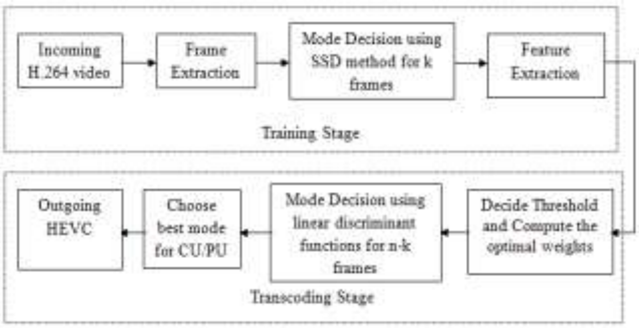

Figure 1: Block diagram of proposed method

A. *Training Stage*

For training K training, two steps should be followed as 1) mode decision using SSD method and 2) feature extraction.
In the first stage of the training, for K frames decision of modes is done by using sum of square difference (SSD) which is the normal method of transcoding. In this method, block division of frame is done of size 64x64. Now DCT transform of that block is taken and quantization is applied on it. Then entropy coding of quantized block is done. Now reversed procedure is carried out as decoding the bitstream, inverse quantization and inverse transform. The sum of square difference is calculated between original block and reconstructed block. Now the block of 64x64 is divided in to four 32x32 blocks and SSD of each block is calculated. If the addition of SSD of four 32x32 block is less than the SSD of original 64x64 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).

During the normal transcoding method i.e. SSD method, different features are extracted which will be used for mode decision in transcoding stage. These features are as follow,

1) Motion vector variance of block
2) Motion vector variance in X direction
3) Motion vector variance in Y direction
4) Energy of DCT coefficient

These features can be used for mode decision because these features have monotonic relationship with SSD. If the value of feature is more than threshold which is calculated in transcoding stage then block must be divided in to sub blocks

B. *Transcoding Stage*

There are two steps in transcoding stage as 1) Decide threshold and compute the optimal weights, and 2) Mode decision using linear discriminant functions for n-k frames.

The MV variance has a very high correlation with one of the two classes i.e. split and unsplit. A block with a high value for the MV variance distance is most likely to be split. For this reason, the incoming blocks with a MV variance distance υ higher than a threshold are removed from the set on the assumption that they shall be split. For the rest of the incoming blocks (i.e., for which $\upsilon \leq$ Tdhigh), the classification is applied using the linear discriminant function solution [2].

The feature value is stored in array X and its mode is stored in array y0 and y1 for split and unsplit for each depth. Then for each feature, optimum weight is calculated using following equation,

$$w_i^{opt} = (X^T X)^{-1} X^T y_i \quad (1)$$

For mode decision, optimum weight is multiplied to the feature vector which is calculated during transcoding stage and score of two modes is calculated using equation,

$$s_i = X_c w_i^{opt} \quad (2)$$

If the score of split mode is greater than score of unsplit mode then block will be divided in four parts and all procedure is repeated for each subdivided blocks.

IV. EXPERIMENTAL RESULTS

The experiments are designed to provide answer to the question that how many frames are needed during the training process to build an efficient model. Thus, three different video sequences are first transcoded using trivial method. Then

ISSN: 2321-8169
4254 – 4256

proposed method is applied on same sequences. To see the effect of variation of number of training, first 25% of frames of sequence are used for training and BD-rate and speed up factor is calculated. Now training frames are increased to 50% and results are calculated. The overall results are given in table II and table III.

TABLE II. TRANSCODER RESULTS USING 25%TRAINING FRAMES

Sequence	Method	BD-Rate(%)	Speed-up
BQ Mall	Trivial	0.0	1.0
	Proposed	4.5740	2.0555
Party Scene	Trivial	0.0	1.0
	Proposed	8.3874	1.3798
Race Horses	Trivial	0.0	1.0
	Proposed	5.3728	1.9849

TABLE III. TRANSCODER RESULTS USING 50%TRAINING FRAMES

Sequence	Method	BD-Rate(%)	Speed-up
BQ Mall	Trivial	0.0	1.0
	Proposed	2.4960	1.5875
Party Scene	Trivial	0.0	1.0
	Proposed	3.9502	1.3477
Race Horses	Trivial	0.0	1.0
	Proposed	2.7839	1.5238

From table II and III, it is clear that, using high number of training frames gives better rate distortion performance. But at same time, speed-up factor reduces. For sequence BQ mall, distortion loss reduces from 4.5740 to 2.4960 as increasing number of training frames from 25% to 50% at cost of reduce in speed of transcoding.

.

V. CONCLUSION

Transcoding solution based on linear discriminant function is proposed in which features are extracted during training stage of n-k frames and these features are used to decide CU mode of HEVC. From results, it is conclude that higher the number of frames used for training, the better the rate distortion performance at cost of reduction in speed-up factor.

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.