Homogenous HEVC video transcoding by transform coefficient removal

Krzysztof Wegner, Damian Karwowski, Krzysztof Klimaszewski, Jakub Stankowski, Olgierd Stankiewicz, Tomasz Grajek
Chair of Multimedia Telecommunications and Microelectronics, Poznan University of Technology, Polanka 3, 60-965 Poznań
dkarwow@multimedia.edu.pl

***Abstract* – In this paper, a novel method of transcoding of HEVC encoded video is presented. The novel approach enables a very fast transcoding without the need for the image reconstruction. It is based on the principle of removing of the carefully selected coefficients from the bitstream. The removal of coefficients is done on the bitstream level and can easily be tailored to software as well as hardware implementations. The technique is able to provide a bitrate reduction in the order of up to 10%, with the image quality decrease of about 0.2-0.5dB.**

***Keywords* – HEVC transcoding; homogenous transcoding; video compression**

I. Introduction

The transmission of the video in heterogeneous telecommunication networks is possible only when the requirements of the consecutive segments of the network are met. One of the most important requirements is the maximum bitrate allowed by the network. For different applications, like mobile TV, digital terrestrial TV or IP transmission, have different limitations in this respect. The most basic approach to solving this problem is to compress the video with different target bitrates and store them separately. This, however, increases significantly the required storage space, which may pose a serious problem due to the costs. A different approach is to store just the best quality bitstream and transcode it on-demand to suit the bitstream to the requirements of the recipient's network. The latter approach becomes even more popular due to reduced storage space requirements and increasing processing capabilities in the form of cloud computing, available on-demand. The problem is, however, the development of the transcoders that are able to perform the processing on-line and still provide a good quality of the reconstructed video.

II. HEVC transcoding

A classical approach to the transcoding involves a very simple concept of the cascaded decoder of the original stream, followed by the encoder that encodes the reconstructed video so that the resulting bitstream meets the requirements that are imposed by the network. The most important requirement is the target bitrate. The encoder can encode the video while optimizing the quality of the transcoded video using the specified budget of data.

This approach, however, is burdened by the most obvious limitation – the full decoding of the video, followed by the full encoding, is very resource-hungry. It requires not only sufficient processing power, but also sufficient memory to buffer the decoded video prior to the encoding and significant amount of memory bandwidth required by encoder and decoder. Also, the second phase – encoding itself – is an extremely complex process and is best avoided in any processing of the video. The complexity of the process and the resources required, make it very difficult to implement such a technique on battery powered mobile devices, and even for stationary installations may be inconvenient.

To remedy that, new transcoding algorithms are developed, that aim to decrease the complexity of the process. They achieve that goal by omitting some steps of decoding of the bitstream or by simplifying the encoding process of the transcoded video. Significant number of publications describe fast transcoding methods for HEVC encoded bitstreams [2-7]. The source of the complexity reduction in those transcoders is the omission or simplification of certain phases of encoding by reusing the information from the original bitstream. Another popular approach is to use the information from the original stream and to exploit the statistical dependencies between the modes used in the original bitstream and the modes used in the transcoded bitstream. This makes it possible to simplify the encoding process while enabling some level of optimization during the encoding. The most popular methods use such an approach for block division selection (CU, PU and TU divisions), motion estimation data (motion vectors, reference frame index/indices), encoding modes (interframe or intraframe). The algorithm that simply reuses the data may be augmented by the information about the statistical information about the division/motion/mode correlation that was mentioned before, thus, a limited set of possible encoding variants are verified, instead of the verification of every possibility. Such methods can provide about 70% of transcoding time reduction, while the encoding efficiency is reduced by about 3% when compared to the classical cascaded transcoder, as calculated by the Bjontegaard metric [8].

III. The proposed HEVC transcoder

A. The main idea

The idea of reusing the data that is employed in the recently developed transcoders leads to a simplified transcoder

The research concluded in this paper were conducted during the execution of the research project LIDER, founded by The National Centre for Research and Development (NCBiR), project number LIDER/023/541/L-4/12/NCBR.

architecture, but still the transcoding is performed in the same way as before – first the original bitstream is decoded, and then encoded using a simplified (due to the reusing of data) encoder. In this work, a different approach is proposed, that does not perform the full reconstruction of the video and does not operate on image samples level. The structure of such an simplified transcoder is presented on Fig. 1.

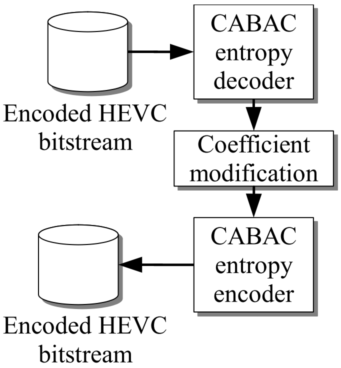

Figure 1. The proposed video transcoder structure – consecutive processing phases

In the first stage, the proposed transcoder performs only the entropy decoding of the bitstream data and just extracts the quantized DCT transform coefficients of the prediction error, that are included in the bitstream. The next step is the modification of those coefficients (described in the following part), and finally, the third part of the process is the entropy encoding of the modified coefficients. Due to such a simplified algorithm, that entirely omits the reconstruction process, the implementation of the algorithm is very efficient using both, software and hardware realizations and does not need bandwidth consuming operations like motion compensated prediction. It must be highlighted here, that the proposed in the paper idea of video transcoding has been already investigated in the context of the older technologies, like AVC/H.264 [10-12]. Because the results obtained for AVC proved to be very promising, the authors decided to explore this transcoding idea in the context of the newest HEVC/H.265 video coding. It must be stressed here, that due to the strong differences between the two technologies (older AVC and the new HEVC) and higher complexity of the HEVC in general, the application of the method in the new codec put the requirement of solving a number of new technical and implementation problems.

B. Quantized DCT transform coefficient modification algorithm

Each of the DCT transform coefficient of the residual data, is characterized by a certain influence on the quality of the image seen by the viewers. At the same time, each of the coefficients has an influence on the size of the encoded bitstream. By analyzing both of those influences, a certain conclusion can be drawn, that leads to the following proposal of transcoding strategy that leads to reduction of bitrate.

It is proposed to remove from the input HEVC bitstream only the coefficients that have the value of ±1. Due to removing of the coefficients of such a low value, the decrease in quality of the reconstructed image is minimal, while still allowing for a certain level of change of the bitrate.

In the proposed method, the decision to remove a certain coefficient is preceded by analysis of the influence of the removal of this coefficient on the prediction signal in the neighboring blocks. This is done because in HEVC, the method and the result of the prediction of the image data is done based on the values of pixels in the neighboring blocks. By changing the coefficient value, on influences the pixel values in the reconstructed image. Thus, a special care must be exercised in order not to introduce significant prediction errors in the neighboring blocks, and not to cause a coding drift. In case when removing of a certain DCT coefficient influences the neighboring blocks in a significant way, the coefficient should not be removed.

Another important case that needs to be considered is when the coefficient that is considered to be removed is the only non-zero coefficient in the transform block. In such a case, the removal of the coefficient would result in the necessity to update some other parts of the bitstream (like CBF – Coded Block Flag or residual quadtree structure) that would cause the need to decode and encode more data. Therefore, to avoid a significant increase in complexity of the process, the removal of such coefficients is not performed in the proposed coder.

After each modification of the contents of the transform unit (TU), it is necessary to perform the process of the hiding of the sign of the coefficients (as described in the MPEG-H HEVC/H.265 standard). In order to do so, it is necessary to check whether specific conditions are met that enable the sign hiding, and if possible, the whole process is performed according to the standard.

C. DCT coefficient removal strategy

The aim of the coefficient removal strategy employed in the proposed transcoder is to minimize the quality loss of the transcoded video, while reducing the bitstream size. In the method proposed, the degree of the bitstream size reduction is managed on two levels: frame level and CTU level.

The procedure of the coefficient removal is preceded by the selection of candidate coefficients for removal. The candidates are then verified according to the previously described rule. The remaining proceed to the next step. The removal strategy assumes that at first the high frequency coefficients are removed from the TUs in the frame. At the TU level, the removal continues until the desired number of coefficients with value ±1 are removed.

IV. Assessment of the Performance of the Proposed Method

A. Methodology of the tests

The implementation of the proposed transcoder is based on the HEVC model software version HM 13.0 [9]. The transcoder software performs consecutive calculation phases, as described in the previous points. Due to the very high complexity of the HEVC codec, the development of the transcoder was very difficult and time-consuming.

In order to perform an experimental verification of the performance of the presented method, a set of encoded HEVC streams was prepared. The streams were obtained by compressing a set of test sequences using the unmodified HM software. This stage of the experiment was conducted in accordance to the recommendations specified by the ISO/IEC and ITU-T in the document called "*Common test conditions*"[1]. Compression of the test sequences (class B and C) was repeated for different settings of QP index value (20, 25, 30, 35, 40, 45, 50), in order to be able to verify the performance of the method in the wide range of bitrates and qualities. The bitstreams were then used as input to the transcoding process. The transcoding was done with certain target bitrate setting.

During the initial coding as well as during the transcoding, a set of data was gathered about the bitrate and the quality of the reconstructed video. Each time, the quality of the reconstructed video was calculated using the PSNR metric, calculated between the original uncompressed sequence and the reconstructed version.

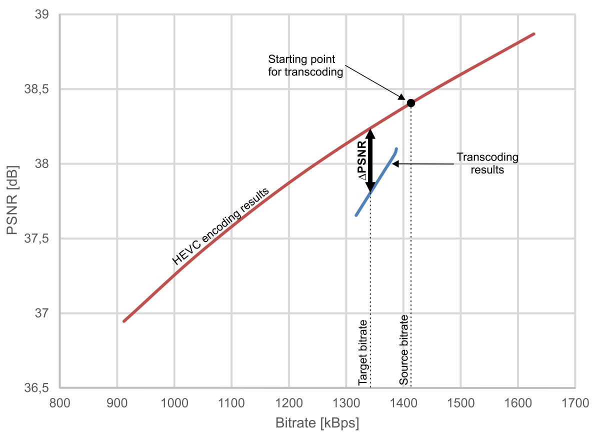

Figure 2. The ΔPSNR value as calculated in the tests

In the next step, a ΔPSNR metric was calculated in the following way. First, the original bitstream was transcoded to a specific target bitrate. The actual bitrate of the obtained stream was noted, and then, the original uncompressed sequence was compressed using a regular HM software with the target bitrate setting equal to the actual value obtained as a result of the transcoding. Then, the PSNR metric was compared for both cases and the difference between those two values was obtained as ΔPSNR. The idea is illustrated on the Fig. 2.

B. *Experimental results*

The results for the whole set of sequences allow to draw some general conclusions, stated below. As the example of the results, the results of the test transcoding for the "Kimono1" sequence are presented on Fig. 3.

The removal of up to a single ±1 DCT coefficient in each TU block leads to an approximately 1% of bitrate reduction. This bitrate reduction results in some 0.25-0.5 dB of quality loss. To obtain a higher reduction of the bitrate, more coefficients need to be removed from the TUs. For the case when up to two coefficients were removed from each TU, the reduction of bitrate reached 2% while the quality loss remains at the level of 0.25-0.5 dB. Removal of up to 16 coefficients in each TU allows the bitrate reduction of about 10%, while the quality loss is about 0.5dB.

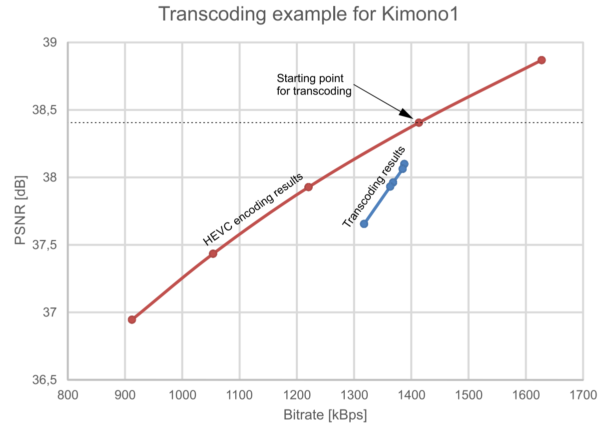

Figure 3. The transcoding results for the "ParkScene" sequence

Aggregated results of the experiments are presented the Tables I to III. For each tested QP value, the average is taken over all tested sequences and extreme values are given as well.

TABLE I. TRANSCODING RESULTS FOR SCENARIO WHEN UP TO 1 COEFFICIENT WAS REMOVED FROM EACH TU

QP	**bitrate reduction [%]**			**ΔPSNR [dB]**		
	average	***min***	***max***	***average***	***min***	***max***
20	2.68	1.40	4.50	0.38	-0.14	1.58
25	2.18	0.90	5.10	0.63	-0.22	3.53
30	1.46	0.60	3.30	0.34	0.03	1.14
35	0.93	0.40	1.70	0.60	0.10	1.28
40	0.66	0.30	1.50	0.42	0.02	2.07
45	0.54	0.20	1.30	0.21	0.04	0.68
50	0.43	0.20	0.80	0.11	0.02	0.21

TABLE II. TRANSCODING RESULTS FOR SCENARIO WHEN UP TO 2 COEFFICIENT WAS REMOVED FROM EACH TU

QP	**bitrate reduction [%]**			**ΔPSNR [dB]**		
	average	***min***	***max***	***average***	***min***	***max***
20	3.14	1.90	5.10	0.48	-0.11	1.71
25	2.68	1.20	6.70	0.31	-0.18	1.13
30	1.78	0.70	4.30	0.34	-0.03	1.16
35	1.13	0.50	1.90	0.31	-0.13	1.18
40	0.79	0.30	1.70	0.45	0.03	2.07
45	0.66	0.30	1.40	0.21	0.05	0.68
50	0.51	0.30	0.90	0.11	0.03	0.21

TABLE III. TRANSCODING RESULTS FOR SCENARIO WHEN UP TO 16 COEFFICIENT WAS REMOVED FROM EACH TU

QP	**bitrate reduction [%]**			**ΔPSNR [dB]**		
	average	***min***	***max***	***average***	***min***	***max***
20	9.24	4.70	17.50	0.92	0.01	2.48

25	6.85	2.60	18.20	0.55	0.03	1.48
30	3.95	1.40	8.70	0.44	0.05	1.10
35	2.34	0.80	5.40	0.48	-0.09	1.25
40	1.59	0.50	4.20	0.52	0.03	2.09
45	1.11	0.40	2.50	0.16	0.06	0.37
50	0.78	0.40	1.30	0.12	0.03	0.22

On the Fig. 4 and Fig. 5, a sample plots are shown, that display the dependency of the bitrate reduction ratio on the number of coefficients removed from each TU for different QP values used for initial coding of the original bitstream. As expected, the reduction is the highest for low QP values, for which the number of non-zero coefficients in the TU is larger than for higher QP values. Even for QP = 50 the bitrate reduction is achieved, despite the fact that for such a high QP value, the number of non-zero coefficients is significantly reduced.

V. CONCLUSIONS

In this paper, a novel concept of homogenous transcoder was presented, that is characterized by a ultra-fast processing time, compared to contemporary transcoding methods. Low computational cost is due to the reduction of the necessary calculations and processing steps. The proposed transcoder operates directly on the components of the data stream, therefore no time-consuming processes, such like the reconstruction of the video, is required. A large performance boost is due to the fact that the proposed transcoder does not perform any part of the video encoding – nor full encoding, nor simplified version of the reconstructed video. The only steps that are necessary are, apart from the actual coefficient removal and correction process, entropy decoding of the original bitstream transform coefficients and encoding of the reduced number of the transform coefficients afterwards.

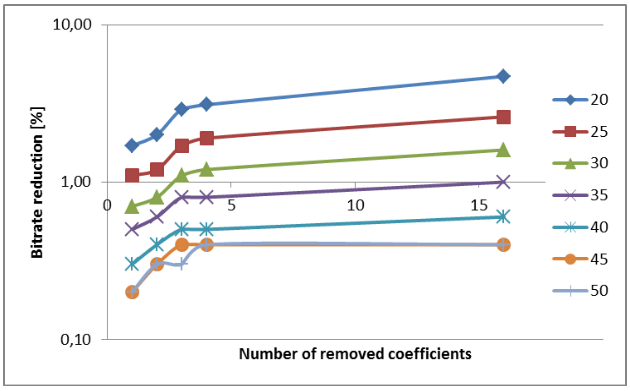

Figure 4. Bitrate reduction ratio for BasketballDrill sequence for different QP values

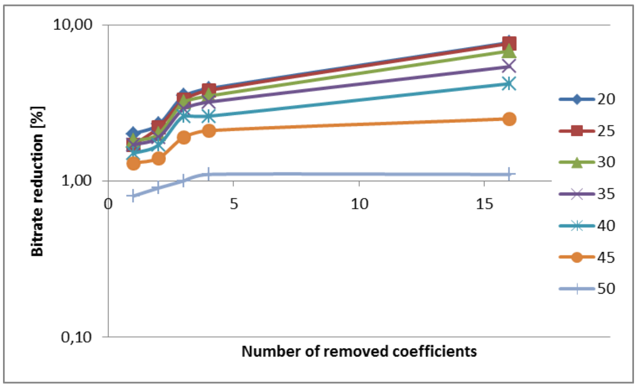

Figure 5. Bitrate reduction ratio for Kimono sequence for different QP values

As a general conclusion from the research. the following can be stated. The reduction of bitrate of a HEVC encoded bitstream by 1%-10% is easily obtained using the proposed transcoder. The average quality loss is meanwhile moderate, at up to about 0.5dB compared to the quality of the video that would be encoded at the same bitrate from the original video, using a regular HEVC coder.

In this paper, only the results for the case when the removal of coefficients is limited to the coefficients with the value of the amplitude equal 1. The same idea can be. however. extended to the removal of the coefficients with greater amplitudes. Also, the number of the removed coefficients can be easily controlled at the frame level, allowing the removal of coefficients mainly from the TUs that are less susceptible to the loss of quality. Another viable method would be to reduce the amplitude, instead of remove the coefficients with larger amplitudes. In some cases this may bring a similar bitrate reduction with lower quality loss.

The possible modifications lead to significantly greater bitrate reduction, but they require a precise and accurate drift control methods. This provides for an interesting area for further research.

REFERENCES

[1] Bossen F., 2012, „Common test conditions and software reference configurations", Joint Collabo-rative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11. Doc. JCTVC-J1100.

[2] L. P. Van, J. D. Cock, G. V. Wallendael, S. V. Leuven, R. Rodriguez-Sanchez, J. L. Martinez, P. Lambert, R. V. Walle, "Fast transrating for high ef-ficiency video coding based on machine learning", IEEE ICIP 2013, 2013.

[3] Luong Pham Van, Johan De Praeter, Glenn Van Wallendael, Sebastian Van Leuven, Jan De Cock, Rik Van de Valle, "Efficient Bitrate Transcoding for High Efficiency Video Coding", IEEE Transactions on Multimedia, Vol. 18. No. 3, March 2016.

[4] L. P. Van, J. D. Cock, A. J. Diaz-Honrubia, G. V. Wallendael, S. V. Leuven, R. V. Walle, "Fast motion estimation for closed-loop HEVC transrating", IEEE ICIP 2014, 2014.

[5] Z. Lin, Q. Zhang, K. Chen, J. Sun, Z. Guo, "Efficient arbitrary ratio downscale transcoding for HEVC", 2016 Visual Communications and Image Processing (VCIP), Chengdu. 2016,

doi: 10.1109/VCIP.2016.7805465.

[6] V. A. Nguyen, M. N. Do, "Efficient coding unit size selection for HEVC downsizing transcoding", 2015 IEEE International Symposium on Circuits and Systems (ISCAS), Lisbon, 2015. pp. 1286-1289.

doi: 10.1109/ISCAS.2015.7168876

[7] L. P. Van, J. De Praeter, G. Van Wallendael, J. De Cock, R. Van de Walle, "Performance analysis of machine learning for arbitrary downsizing of pre-encoded HEVC video", IEEE Transactions on Consumer Electronics. vol. 61, no. 4. pp. 507-515, November 2015,

doi: 10.1109/TCE.2015.7389806

[8] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves", ITU-T SG16, Doc. VCEG-M33, Apr. 2001.

[9] Software model HEVC codec, available online on-line at https://hevc.hhi.fraunhofer.de/svn

[10] M. Domański, J. Marek, Fine Grain Scalability of Bitrate Using AVC/H.264 Bitstream Truncation, 27th Picture Coding Symposium, PCS 2009, Chicago, USA, 6-8 May 2009, pp. 1-4. (www.multimedia.edu.pl)

[11] J. Marek, M. Domański, Direct Transcoding of AVC/H.264 bitstreams for bitrate reductions with minimized video quality losses, The 2009 International Conference on Image Processing, Computer Vision and Pattern Recognition, Las Vegas, USA, 13-16 July 2009. (www.multimedia.edu.pl)

[12] J. Marek, Homogenous Video Transcoding of H.264/AVC Intra Coded Frames, International Conference on Computer Vision and Graphics, Warsaw, Poland, 10-12 November 2008. (www.multimedia.edu.pl)