INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG2016/M39528
October 2016, Chengdu, China

Source	**Poznań University of Technology**
Status	**Input**
Title	**Pruning of insignificant coefficient in HEVC transcoding**
Author	**Krzysztof Wegner, Tomasz Grajek, Jakub Stankowski, Damian Karwowski, Krzysztof Klimaszewski, Olgierd Stankiewicz, Marek Domański**

1. Introduction

Heterogeneous network environment often requires the use of different bitrate of a video stream when transmitting over different part of network. One of the most popular solutions is to transcode the original video stream to a new bitstream at the required bitrate (different from the original one).

This document refers to homogenous HEVC video transcoding, that transrates the source HEVC bitstream by removing (setting to zero) some of the higher frequency transform coefficient(s) in the Transform Unit (TU) blocks. The document presents technical details of the proposed HEVC transcoder together with a set of experimental results showing its performance and computational complexity. The aim of this document is to bring attention of MPEG experts to real possibilities of fast transrating of HEVC encoded bitstreams, and to give numerical data reflecting the complexity of the task and the expected quality loses and drift.

2. HEVC Video Transcoding – mainly used solutions

Classical solution for a video data transcoding is to perform full decoding of source bitstream followed by the re-encoding of earlier decoded images with new set of parameters like bitrate, qp value or the resolution. The above-mentioned approach is commonly known in the literature as cascaded pixel domain video transcoder (CPDT). CPDT actually shows a high computational complexity, however, since this is the most straightforward approach to transcoding, a cascaded transcoder is currently the subject of most research works.

In the CPDT, the trivial connection of a video decoder with a fully separate video encoder leads to a high complexity of the entire system. For this reason, the main research direction aims at reducing transcoding complexity. The source video bitstream and the transcoded one share many properties, mainly because both of them are representations of the same video. Both can have similar image partitioning onto Coding Units (CUs) or they can have similar motion vectors as well as any other information used to represent video content. Very popular way to reduce complexity of the CPDT is the exploit the similarity between source video bitstream and the generated transcoded one, by guiding the encoder with the information from the source bitstream. Such an approach (called in this document a guided transcoding) provides measurable benefits, because some of information extracted by the decoder can simplify significantly the procedure of mode selection in the encoder. Good examples of homogenous HEVC transcoders that are based on this idea are [2, 3, 4]. In the cited proposals the basis for reduction the complexity of a video transcoding is:

- High similarity of a way of partitioning of images into CUs and PUs in both the source and the transcoded data streams [2, 4].
- Similarity of motion information (e.g. motion vectors, reference frames indices) in both data streams [3].

Although guided transcoding leads to solutions of lower complexity when compared to trivial CPDT (even 80% complexity reduction under 3% BD-Rate increase can be achieved), there is still a need of full decoding and reconstruction of a source video followed by re-encoding and selection of compression modes for the decoded video.

3. Performance of the CPDT Transcoding

Transcoding a video always introduces some loss of video quality. This is also the case of CPDT based on HEVC technology. In order to evaluate the quality loss we did the following experiment. At first a set of 1920x1080 test video sequences has been encoded (with the use of JM 13.0 software) with all QP values from 0 to 51 resulting in 52 source bitstreams per sequence. Then, each resultant bitstream was fully decoded and then re-encoded with all QP values again ranging from 0 to 51. This gives 52 target bitstreams per a source bitstream, and 52^2 target bitstreams per sequence. Each target bitstream was decoded again and the quality of the resultant video was measured by PSNR (with respect to original video).

The quality loss (ΔPSNR) of transcoded video has been expressed as the difference between PSNR of video decoded from target bitstream and the PSNR of video decoded from source bitstream, under the condition that both – the source and the target bitstream have the same bitrate. In other words, we compared the video quality difference between the two cases. First, when the video bitstream was originally encoded at requested bitrate directly from the original data, and the second, when the video bitstream at requested bitrate was created based on the source bitstream with higher bitrate in a process of transcoding. The ratio between the source bitstream rate and the requested bitrate is called transcoding ratio. The observed quality loses of the video decoded from second bitstream is directly caused by the transcoding process. Of course the higher the bitrate of the source bitstream used, the better the quality of the transcoded video. A question arises: is there any relationship between transcoding ratio and the quality losses caused by the transcoder?

Fig 1. Presents average results (over many test sequences) of the quality losses versus transcoding ratio

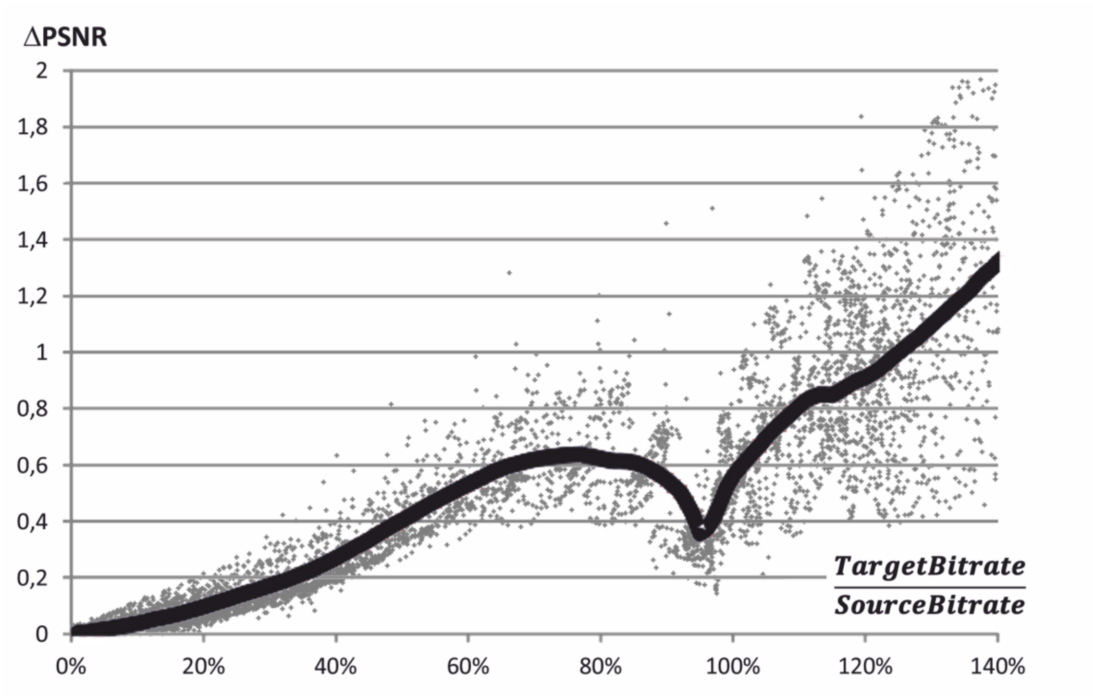

Fig. 1. ΔPSNR (quality loss) with respect to transcoding ratio (i.e. *TargetBitrate*/*SourceBitrate* ratio). The result of averaging the partial data obtained for a set of 1920x1080 test sequences and different values of *SourceBitrate*.

As it can be seen, if one transcodes the video at exactly the same bitrate (transcoding ratio equals 100%), the transcoding causes 0,6 dB quality loses just by quantizing the video content again. The quality loss gets lower if the transcoding ratio deceases, but then rises again at about transcoding ratio of 97%. Between transcoding ratio of 60% and 85% the quality loses are the highest - on average 0,63 dB.

4. Proposed Homogenous Transcoder

When compared to CPDT, transrating the HEVC encoded bitstream can be done much easier, with less computational power, especially when considered in the scenario of a moderate reduction of bitrate.

The main content of the bitstream are code-words of quantized transform coefficients [5]. Also the amount of residual signal in a form of quantized transform coefficients relates directly to video quality. Sending less transform coefficients degrades video quality but also lowers the required bitrate. Commonly, quantization is a way to reduce the amount of data to be transmitted. Unfortunately, quantization, or precisely speaking - requantization done in the CPDT, is the main cause of the rapid quality losses in transcoding.

There is yet another way to steer bitrates, used for instant in JPEG 2000. Instead of further quantize the coefficients, we can transmit only part of them, skipping less important ones.

Exactly this idea is a key to the proposed fast video transcoder.

Instead of decoding and reconstructing video samples, and then re-encoding them again, we simply remove less significant quantized transform coefficient from the source bitstream, and thus create the target one. We remove as many transform coefficient as it is required to meet the target bitrate.

In the presented results we start with the least significant transform coefficients (high frequency ones), that have been quantized to 1. Such coefficients do not contribute much to the overall image quality, and their removal will not degrade an image much.

It can be noted, that the described method is not restricted to removal of those particular coefficients, but in principle, any transform coefficient, even those with larger values, can be removed, as long as its removal will not contribute much to image degradation.

4.1. Processing pipeline

In the proposed fast transcoding, we try to avoid as many coding/decoding phases as possible. In the proposed approach only the following actions are performed:

1. The input bitstream is analyzed and the entropy decoding is performed on the source data stream. Since in HEVC there is no possibility to perform partial entropy decoding, the whole data stream has to be decoded. The goal of this step is to decode quantized transform coefficients.
2. The decoded quantized transform coefficients are scanned to list all candidate coefficients for removal. The candidate selection process and the removal process is described in the following paragraph.
3. Selected quantized transform coefficients are being removed. Their value is set to zero, as the HEVC is not transmitting zero coefficients.
4. Next, all TUs in which some coefficients have been removed are adjusted to properly signal the hidden sign.
5. The CBF fields of the modified TUs are adjusted.
6. Finally, all syntax elements are entropy encoded. This concludes the transcoding process in our approach.

The described process can be repeated many times over the same set of data, and the transcoding errors will not be greater than in the case of a single pass transcoding which removes the same coefficients. Therefore, instead of transcoding in a single pass and removing many coefficients, it is possible to remove e.g. half of them, and then, depending on whether the target bitrate has already been achieved or not, decide on removing the other half or not in next pass.

4.2. Transform coefficient removing strategy

The transform coefficient removal strategy assumes that the removal of a coefficient should cause as little image quality degradation as possible, while decreasing the bitrate.
The coefficients are marked as candidates for removal on the TU level, starting from the highest frequency coefficients and moving gradually towards the DC component, until the desired number of coefficients are removed for a given TU. At this stage, only the coefficients with value equal to 1 are removed.

To preserve the image quality, a certain restriction has to be applied to the process:

- The TU of Intra CUs are excluded from the removal process, as the modification of the residual signal could change the intra prediction signal. Such drift accumulates very fast throughout the image causing visible artifacts.
- Similarly, any TU neighboring any TU of an Intra coded CUs are being excluded. This is done in order to preserve the quality of Intra coded blocks, since their quality

is important in the inter prediction process of further frames and modifications of coefficients would cause drift and excessive loss of image quality over time.

4.3. Drift in the proposed approach

Of course, our approach introduces a certain amount of drift to the encoded sequence, since the residual signal changes while the prediction modes remains unchanged and thus cannot adapt to the modified data in the CUs that went through the process of coefficient removal. However, with careful selection of the coefficients from TUs for the removal process (as described in the previous point), we are able to limit the amount of drift to acceptable levels for most of the cases.

5. Methodology of the experiments

For the test purposes we have implemented our method on top of HTM version 13.0. The prepared software enables the following:

1. Decoding of all syntax elements which are contained in the HEVC data stream without decompressing them and reconstructing the samples.
2. Analysis of the decoded quantized transform coefficients and removal of the selected coefficients. Once the selected coefficients are removed, the software adjusts the values of some other syntax elements if needed (e.g. value of CBF flag must be changed when removing all non-zero transform coefficients in a TU block).
3. Entropy re-encoding of the modified set of syntax elements.

Parameters of the proposed transcoder have been evaluated in a series of experiments. The goal was to assess the possibility of bitrate reduction and the loss of video quality when transrating HEVC encoded bitstreams under the assumed scenario of removing transform coefficients. The experiments were performed according to "common test conditions" (CTC) [1]. Both B and C classes of test video sequences were used, which were encoded with the HM 13.0 reference software configured according to CTC. The encoding of sequences has been performed for values of QP = 20, 25, 30, 35, 40, 45, 50. The encoded data streams, obtained in this way, were then fed to a transcoder, who did transrating of streams.

For the first cycle of video encoding, as well as in the case of the encoded bitstream transcoding we report bitrate of the encoded data streams together with quality of the encoded material. In each case, quality of the encoded video has been expressed as a PSNR calculated between the original and reconstructed video.

6. Performance of the proposed transcoder – results

As can be seen from the results of Table 1 and Table 2, removing up to one transform coefficient of value 1 (in each block of TU) leads to a small reduction of bitrate (1% in average) with an average loss of video by 0.25 dB and 0.51 dB for B and C classes respectively.

Table 1. Results of bitstreams transcoding in the scenario of removing up to one transform coefficient of value equal to 1 within each of TU. Results for the B class of test sequences.

		Class B				Transcoded (removing up to 1 coeff)					
	QP	Bitrate [kbps]	PSNR Y	PSNR U	PSNR V	Bitrate [kbps]	PSNR Y	PSNR U	PSNR V	ΔPSNR	$\frac{target\ bitrate}{source\ bitrate}$
Kimono1	20	6697	42.20	43.96	45.96	6565	42.15	43.91	46.01	0.05	0.98
	25	2835	40.55	42.63	44.10	2786	39.35	42.45	44.21	1.20	0.98
	30	1380	38.41	41.45	42.66	1355	38.10	41.45	42.66	0.31	0.98
	35	683	36.00	40.31	41.49	672	35.45	40.31	41.49	0.55	0.98
	40	343	33.59	39.57	40.82	338	33.34	39.46	41.04	0.25	0.99
	45	162	31.20	38.57	39.97	160	30.89	38.57	39.97	0.30	0.99
	50	72	28.95	37.06	38.65	72	28.84	37.06	38.65	0.11	1.00
ParkScene	20	10451	40.99	43.05	44.66	10052	41.04	43.12	45.01	-0.04	0.96
	25	4296	38.56	41.24	42.37	4184	38.45	41.21	42.63	0.11	0.97
	30	1956	35.98	39.66	40.70	1922	35.64	39.53	40.67	0.33	0.98
	35	894	33.39	38.07	39.36	883	32.07	37.68	39.24	1.32	0.99
	40	406	30.95	37.10	38.69	402	28.85	37.12	38.89	2.10	0.99
	45	170	28.70	36.01	37.99	168	28.53	36.01	37.99	0.17	0.99
	50	67	26.76	34.66	36.95	67	26.66	34.66	36.95	0.10	1.00
Cactus	20	33861	39.31	40.67	44.25	32834	40.77	41.62	44.64	-1.46	0.97
	25	8721	37.52	39.43	42.60	8582	37.33	39.31	42.42	0.19	0.98
	30	3766	35.80	38.68	41.13	3731	35.61	38.56	40.96	0.19	0.99
	35	1888	33.71	37.81	39.62	1875	33.56	37.81	39.62	0.15	0.99
	40	979	31.44	37.13	38.50	973	31.14	37.04	38.36	0.30	0.99
	45	481	29.05	36.17	36.99	478	28.93	36.17	36.99	0.12	0.99
	50	219	26.75	34.67	34.73	219	26.67	34.67	34.73	0.08	1.00
BQTerrace	20	70449	39.05	42.60	44.70	68100	38.72	42.84	44.85	0.33	0.97
	25	14691	35.93	41.36	43.62	13939	36.09	41.29	43.29	-0.16	0.95
	30	4101	34.50	40.29	42.70	3965	34.15	39.98	42.10	0.36	0.97
	35	1706	32.82	39.04	41.57	1684	32.55	39.04	41.57	0.28	0.99
	40	822	30.75	38.23	40.81	816	30.56	38.23	40.81	0.19	0.99
	45	395	28.35	37.27	39.85	393	28.29	37.65	40.13	0.06	0.99
	50	179	25.80	35.90	38.31	178	25.58	35.90	38.31	0.22	0.99
BasketballDrive	20	29927	40.05	44.20	45.67	29505	40.00	44.39	45.83	0.05	0.99
	25	8561	38.17	43.05	43.82	8486	38.24	43.29	43.97	-0.08	0.99
	30	3774	36.46	41.92	42.06	3750	35.99	41.69	41.68	0.47	0.99
	35	1905	34.51	40.72	40.42	1895	34.33	40.69	40.18	0.19	0.99
	40	1025	32.42	39.85	39.24	1020	32.31	39.88	39.26	0.11	1.00
	45	522	30.13	38.51	37.36	520	29.86	38.51	37.36	0.27	1.00
	50	254	27.74	36.53	35.05	253	27.53	36.53	35.05	0.21	1.00
									Average:	**0.25**	**0.99**

Table 2. Results of bitstreams transcoding in the scenario of removing up to one transform coefficient of value equal to 1 within each of TU. Results for the C class of test sequences.

		Class C				Transcoded (removing up to 1 coeff)					
	QP	Bitrate [kbps]	PSNR Y	PSNR U	PSNR V	Bitrate [kbps]	PSNR Y	PSNR U	PSNR V	ΔPSNR	$\frac{target\ bitrate}{source\ bitrate}$
RaceHorses	20	6587	40.51	42.34	43.74	6464	40.16	42.34	43.74	0.35	0.98
	25	2741	37.09	39.92	41.49	2705	33.51	38.57	39.66	3.58	0.99
	30	1252	34.15	38.06	39.73	1244	34.07	38.06	39.73	0.08	0.99
	35	603	31.40	36.38	38.15	600	30.34	36.65	37.69	1.07	1.00
	40	288	28.84	35.32	37.07	287	28.81	35.32	37.07	0.04	1.00
	45	127	26.67	33.82	35.33	127	26.59	33.82	35.33	0.07	1.00
	50	58	24.87	31.91	33.10	58	24.83	31.91	33.10	0.04	1.00
BQMall	20	5196	41.10	44.25	45.93	5052	39.46	43.53	44.97	1.65	0.97
	25	2357	38.79	42.47	43.76	2318	38.57	42.47	43.76	0.22	0.98
	30	1192	36.15	40.87	41.88	1181	36.08	41.04	41.77	0.07	0.99
	35	626	33.39	39.17	40.02	622	33.26	39.17	40.02	0.13	0.99
	40	335	30.64	38.03	38.80	334	29.88	37.89	38.61	0.76	1.00
	45	170	27.89	36.61	37.16	169	27.77	36.61	37.16	0.12	0.99
	50	80	25.36	34.59	34.94	80	25.17	34.59	34.94	0.19	1.00
PartyScene	20	9368	39.75	42.52	43.64	8951	38.64	42.21	42.95	1.11	0.96
	25	4336	36.16	39.88	40.86	4204	35.63	39.88	40.86	0.54	0.97
	30	2090	32.92	37.95	38.80	2048	32.66	37.95	38.80	0.26	0.98
	35	1012	29.91	36.23	36.97	999	28.68	35.80	35.98	1.24	0.99
	40	472	27.04	35.10	35.77	469	26.99	35.10	35.77	0.05	0.99
	45	191	24.37	33.73	34.33	190	23.68	33.49	33.52	0.69	0.99
	50	66	22.17	31.86	32.45	65	22.13	31.86	32.45	0.03	0.98
BasketballDrill	20	4692	41.71	43.96	44.71	4611	41.46	43.99	44.73	0.25	0.98
	25	2270	38.64	41.62	41.91	2246	38.53	41.67	41.96	0.11	0.99
	30	1126	35.62	39.60	39.64	1118	34.44	39.54	39.50	1.17	0.99
	35	574	32.92	37.71	37.55	571	32.63	37.71	37.55	0.29	0.99
	40	313	30.50	36.38	36.08	312	30.42	36.38	36.08	0.08	1.00
	45	160	28.07	34.64	34.11	159	27.95	34.64	34.11	0.12	0.99
	50	75	25.63	32.39	31.55	75	25.57	32.39	31.55	0.06	1.00
									Average:	**0.51**	**0.99**

For comparison purposes one can refer data presented here to the results of CPDT transcoder shown in section 3.

7. Conclusions

The new type of homogenous HEVC video transcoder has been presented, whose idea differs significantly from the CPDT transcoders considered in the literature. Only three stages of calculations are performed in the proposed transcoder:

1. Entropy decoding of syntax elements contained in the encoded data stream.
2. Analysis of the decoded transform coefficients and removal of selected ones. Control of the correctness of values of all the syntax elements.
3. Entropy coding of the modified set of syntax elements.

The proposed transcoder allows for efficient transrating of the encoded data stream, when used in the scenario of a moderate bitrate reduction. This can be realized under significantly lower amount of calculations, when compared to the CPDT transcoders.

Acknowledgement

Research project was supported by The National Centre for Research and Development, POLAND, Grant no. LIDER/023/541/L-4/12/NCBR/2013

8. References

[1] Bossen F.. 2012. „Common test conditions and software reference configurations". Joint Collaborative 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 efficiency video coding based on machine learning", IEEE ICIP 2013, 2013.

[3] 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.

[4] 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.

[5] J. Stankowski, D. Karwowski, T. Grajek, K. Wegner, J. Siast, K. Klimaszewski, O. Stankiewicz, M. Domański: „Analysis of compressed data stream content in HEVC video encoder", International Journal of Electronics and Telecommunications, vol. 61, no. 2, pp. 121–127, DOI: 10.1515/eletel-2015-0015 (2015),