IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received May 24, 2019, accepted June 22, 2019, date of publication July 12, 2019, date of current version August 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2928423

Analysis of Video Quality Losses in Homogeneous HEVC Video Transcoding

TOMASZ GRAJEK, (Senior Member, IEEE), JAKUB STANKOWSKI, DAMIAN KARWOWSKI, KRZYSZTOF KLIMASZEWSKI, OLGIERD STANKIEWICZ, (Member, IEEE), AND KRZYSZTOF WEGNER, (Senior Member, IEEE)

Chair of Multimedia Telecommunications and Microelectronics, Poznań University of Technology, 60-965 Poznań, Poland

Corresponding author: Tomasz Grajek (tomasz.grajek@put.poznan.pl)

This work was supported by the National Centre for Research and Development, Poland, under Grant LIDER/023/541/L-4/12/NCBR/2013.

ABSTRACT This paper presents a detailed, quantitative analysis of video quality losses in a homogeneous HEVC video transcoder together with the analysis of the origin of those quality losses and the influence of quantization step alignment on transcoding. With the use of HM15.0 reference software and a set of test video sequences, the cascaded pixel domain video transcoder (CPDT) concept has been used to gather all the necessary data needed for the analysis. This experiment was performed for a wide range of source and target bitrates. The essential result of the work is an extensive evaluation of CPDT, commonly used as a reference in works on effective video transcoding. So far, no such extensively performed study has been made available. The quality degradation between a transcoded video and a video that would be the result of direct compression of the original video at the same bitrate as the transcoded one has been reported. The dependences between quality degradation caused by transcoding and bitrate changes of the transcoded data stream are clearly presented in graphs.

INDEX TERMS Cascaded pixel domain video transcoder, CPDT, HEVC, H.265, video compression, video transcoding.

I. INTRODUCTION

Techniques of hybrid compression of digital video are the most common methods of video transmission used in contemporary IT networks. The hybrid technique that has been developed over the years, has become the subject of numerous standardization projects of expert groups of ISO/IEC MPEG and ITU-T VCEG (e.g. MPEG-2 [1], H.263 [2], AVC [3], [4], HEVC [5], [6] standards). The constantly changing requirements imposed on video compression algorithms, as well as the improving capabilities of hardware, have resulted in faster and faster changing generations of video compression technology.

Today, the state-of-the-art in the field of video compression is the High Efficiency Video Coding (HEVC) technology [6] that has been jointly developed by ISO/IEC and ITU-T and published in 2013 simultaneously as an international standard of ISO/IEC MPEG-H part 2 and recommendation ITU-T H.265 [5]. The new technique is the successor of the widely used and extremely successful Advanced Video Coding (AVC) technology [4], [7]. Comparing these two techniques, HEVC allows up to a 2-fold reduction in the size of the encoded images without compromising the quality of video [8], and, more importantly, it supports the compression of ultra-high definition video, which is believed to be the future of video systems. For this reason, it is highly expected that within the next few years, the new HEVC technique will replace the currently used AVC technology in nearly all fields.

Similarly to the previous video compression standards (e.g. MPEG-2, AVC), HEVC has been developed with many different applications in mind, i.e., not only digital television but also mobile terminals, in which a codec has to run on devices with significantly different computing capabilities. The requirements differ for each of those applications and environments, therefore, the set of compression tools used also varies, accordingly. Very often, not only the set of tools is limited but also the parameters of the encoded video itself changed. For example, the target bitrate or resolution of a video must be suitable for the given application and target device. That is why different versions of the same

The associate editor coordinating the review of this manuscript and approving it for publication was Pradeep Kumar Gupta.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

video sequence with different parameters – resolution, bitrate, frame rate, etc. are needed. Taking into account that once a sequence is encoded, the original video signal is no longer available (thus, there is no possibility to repeat the process of encoding from the original signal), the only way to produce a required bitstream with new parameters is to use a video transcoder. Because in the considered case of video transcoding only some parameters of the encoded video are modified, and the general coding technology remains the same (in our case, it is HEVC), this type of transcoding is called homogeneous video transcoding. It is important to note here that during video transcoding generally, it is only possible to obtain a bitstream representing the video signal whose parameters, like image quality, are always lower than the parameters of a bitstream before transcoding. One of the goals of the research on transcoding is to minimize the loss of quality caused by re-encoding (transcoding) of a video in a case when the previously lossy compressed video is used to create a new bitstream at a different target bitrate. Especially, the quality (losses) of thus created video in comparison with the quality of a video encoded from the original sequence is an interesting issue.

It is commonly known that in the case of hybrid video compression transcoding always introduces some video quality loss, regardless of the choice of the target bitrate of a transcoded video [9]–[13]. This is a direct result of lossy operations performed each time by a video encoder. Although research on the impact of video transcoding on the quality of video, as well as its impact on the size of the encoded bitstream has already been carried out for the older compression techniques (e.g., MPEG-2 or AVC) [9], [12]–[16], no extensive study of video transcoding performance has been presented in the literature so far. Moreover, the conclusions of the existing partial studies cannot be directly transposed to the new HEVC technology. The main reasons are the differences between the older and the new compression techniques that strongly affect the compression process. A kind of exception here is the work [11] (published by some of the authors of this text), which shows the results of the accumulation of video quality losses in multiple HEVC encoding and decoding cycles. However, the above mentioned work also does not analyze in detail the mechanism of occurrence of video quality losses in the process of HEVC transcoding. Such an analysis is the subject of this work.

II. SPECIFIC OBJECTIVES OF THE WORK

Homogeneous HEVC transcoding has been the subject of research works for a few years now [17]–[25]. The main goal of the already published works was to reduce the computational complexity of HEVC re-encoding by exploiting the high similarity of selected parameters of the original (input) and the transcoded bitstreams. In these works, the performance (computational complexity and coding efficiency) of the proposed solutions in relation to the performance of the reference pixel-domain cascaded transcoder has been compared. The phenomenon of video quality degradation in the transcoder has not been the subject of a detailed analysis. For this reason, the authors of those works did not indicate the optimal conditions for HEVC video transcoding for which the loss of video quality is minimal.

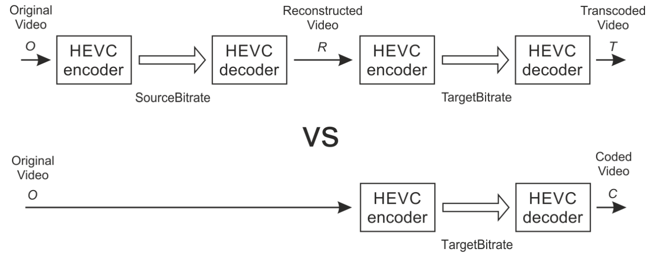

FIGURE 1. Methodology of comparing quality losses during video transcoding.

The goal of this work is to perform a deep study of the impact of video transcoding on the quality of transcoded video material using the newest video compression standard. In this work, a cascaded pixel domain HEVC video transcoder is considered. That is, the process of transcoding is based on feeding a once HEVC-encoded and decoded video to another HEVC encoder (Fig. 1) and performing the full process of compression. As it was mentioned before, such a configuration is commonly used in the literature as a reference in research on the improvement of video transcoding. Therefore, our work may be used as a reference for any future improvements of the transcoding process for HEVC data. The cascaded method can also be used in practice as the simplest possible way of transcoding.

During the work described in this article the authors tried to answer two fundamental questions. First, what is the difference between the quality of an HEVC-transcoded video (transcoded with a specific value of target bitrate) and the quality of a video that would be achieved after a direct compression of the original video (assuming that such a video could be made available) at the same target bitrate? What factors affect this change of quality? Secondly, is there some optimal scenario of video transcoding (understood as a range of changes of bitrate in the transcoder), for which the loss of video quality is minimal? It must be stressed again that in this work ''the loss of video quality'' is expressed as the difference between the quality of a transcoded video (transcoded at some value of a target bitrate) and the quality of a video achieved after encoding the original sequence at the same target bitrate. In order to better understand the mechanism of quality losses, short remarks on this topic are highlighted in the next section.

III. LOSS OF QUALITY IN THE HEVC ENCODER – SHORT REMARKS

A detailed, quantitative analysis of the sources of video quality losses in the HEVC video encoder / transcoder is beyond the scope of this work. Therefore, only some remarks clearing out the reasons of such losses are outlined in this section.

The degradation of quality during transcoding is owed to the fact that generally HEVC is a lossy compression standard, i.e., some data are intentionally omitted in the process of encoding. There are three separate sources of those losses (quality degradation) in the HEVC encoder. The first one is Discrete Cosine Transform (DCT)-like transformation of residual data, which is not fully reversible. The second is the quantization of the resultant transform coefficients. The third and the final source of losses is associated with the modification of reconstructed image samples by in-loop deblocking and sample-adaptive offset (SAO) filters. They aim to improve the visual (subjective) quality, but at the same time, they may introduce value shifts to pixel values, and thus decrease the objective quality of the image.

HEVC exploits finite precision approximation of DCT transformation [4]. Taking into consideration both the coding efficiency and complexity aspects, the HEVC standard requires 16-bit depth of values obtained at each stage of the transform process, including the sign bit (for 8-bit input image sample representation). In order to fulfill this requirement, there is a necessity of bit shifting and clipping operations at some stages of residual data transformation. As a result, a combination of forward (performed in the encoder) and inverse (calculated in the decoder) DCT-like transformations gives only near-lossless reconstruction of input image samples, and not fully lossless. So, there is some (in practice very minor) loss of quality. It should be noted, however, that this quality loss happens twice in the transcoding procedure, when a sample goes through the DCT-like transformation chain: first, during the encoding of the original material, and for the second time, during the transcoding itself.

The step that introduces losses in all hybrid video codecs is the quantization of transform coefficients of residual data [6], HEVC not being an exception from this rule. The process of quantization is equivalent to (however, from the mathematical point of view, not exactly the same as) dividing transform coefficients by the quantization step size (so called *QStep*) together with subsequent rounding operations. This leads to an irreversible loss of part of information about the signal – the reconstructed images differ significantly from the original ones. The level of quality loss depends directly on the quantization step size (the value of *QStep*) – the higher the value, the coarser the representation of transformed residual data. Coarse representation results in a lower quality of reconstructed images, but gives a higher compression ratio of the video.

Due to the nature of block-based encoding used in the HEVC coder, the reconstructed image can exhibit some distortions, usually called block artifacts, which are very annoying for viewers, especially at lower bitrates. In order to improve the subjective quality of the reconstructed images, deblocking and SAO filters are used [6], [26], [27]. This helps to improve the subjective quality, but, at the same time, can degrade the objective quality metrics, such as still commonly used PSNR.

In order to reduce the blocking artifacts (by the use of a deblocking filter) and attenuate the ringing artifacts (with the help of a SAO filter) that are present in the reconstructed images after decoding, the values of image samples are appropriately modified before writing them to a file or showing them on a screen. These filtering operations, when carried out several times (as is in transcoding), may change the values of image samples essentially. Even though the goal of those filters is to improve the subjective quality, the effect of these operations on the PSNR metric (used to measure the quality in transcoding) is difficult to predict and the use of those filters may lead to a further reduction of the objective quality metric value.

IV. LOSS OF QUALITY IN HOMOGENEOUS HEVC TRANSCODING – METHODOLOGY OF EXPERIMENT

Suppose that we have an HEVC-encoded video material, compressed at *SourceBitrate*. We can transcode the bitstream by decoding it and re-encoding at a different *TargetBitrate*, as seen in Fig. 1. The question is: what is the difference between the quality of a transcoded video T (obtained by re-encoding a previously encoded video R to a new bitstream with a rate equal to the *TargetBitrate*) and the quality of a video C achieved by a direct encoding of the original video to a bitstream whose rate is again equal to the *TargetBitrate*? So, what loss of video quality is actually introduced by using the previously encoded video R instead of the original one O. What loss of quality is introduced by video transcoding?

In order to answer this question, an extensive and detailed experiment was conducted. At first, original video materials O have been encoded by an HEVC encoder with the use of all possible quantization parameter values QP (QP ranging from 0 to 51) resulting in a number of bitstreams with different *SourceBitrates*. The resultant bitstreams have been decoded and the obtained video R has been encoded again by an HEVC reference encoder in order to simulate a cascade pixel domain transcoder. Re-encoding of a single decoded video R has been performed with the use of all possible quantization parameters (QP ranging from 0 to 51) in order to simulate a wide range of requested bitrate changes. As a result, more than 2500 video bitstreams have been prepared.

The quality of each decoded video, both after the first encoding (video R) and after transcoding (video T), has been measured by the luminance PSNR metric. Also, the bitrates of each encoded stream have been gathered for further analysis.

The reference HEVC software has been configured according to the "main_randomaccess" test conditions [28], [29] with the in-loop deblocking and samples adaptive offset (SAO) filters turned ON. Such a scenario of video encoding is well recognized as a high-efficiency video compression scenario and corresponds to HEVC encoder configuration which is very commonly used in practice. The widely used *BlueSky*, *PedestrianArea*, *Riverbed*, *RushHour*, *Station*2, *Sunflower* full HD test sequences were taken as the test material (1920x1080 spatial resolution, 30 frames per second). In each case, only the first 200 frames from each sequence were encoded.

In order to determine the loss of image quality that results from cascaded video transcoding, ΔPSNR was calculated as the difference between the PSNR of transcoded video T and

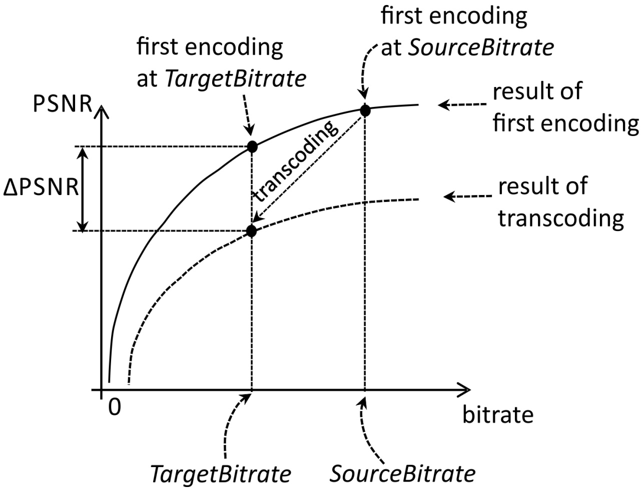

FIGURE 2. The way of determining ΔPSNR in a homogeneous HEVC transcoder.

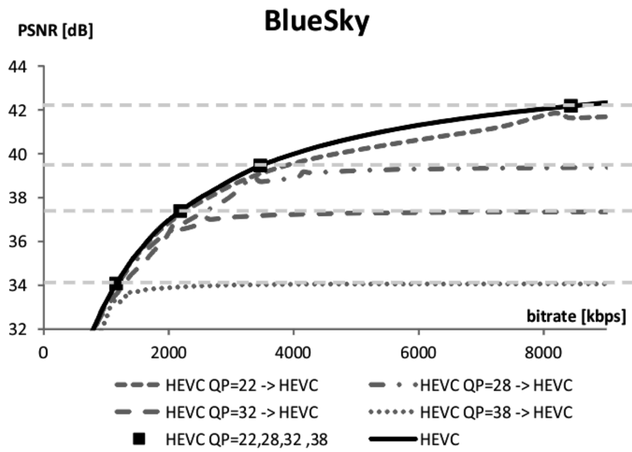

FIGURE 3. Exemplary results of video transcoding obtained for *BlueSky* test sequence.

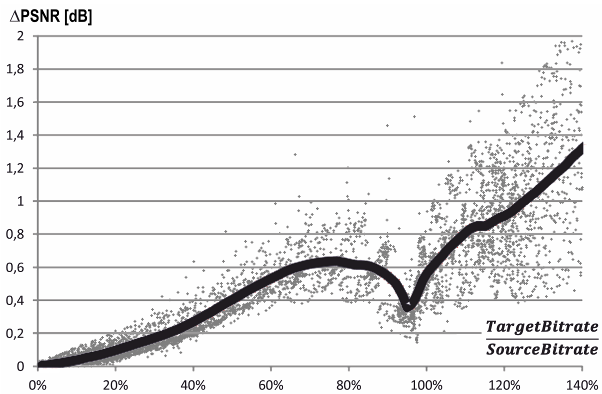

FIGURE 4. ΔPSNR (quality loss) with respect to transcoding ratio (i.e. *TargetBitrate*/*SourceBitrate* ratio). The result of averaging the partial data obtained for individual test sequences and different values of *SourceBitrate*.

the PSNR of encoded video C, both encoded at the same *TargetBitrate* (Fig. 1). The procedure of determining ΔPSNR is also illustrated in Fig. 2.

In case of each transcoding, the relationship between the *SourceBitrate* (i.e., bitrate of the first video encoding) and the *TargetBitrate* (i.e., final bitrate of transcoded data) was additionally analyzed (see Fig. 2). The obtained partial results were used to determine the ΔPSNR in the function of the transcoding ratio (i.e., *TargetBitrate*/*SourceBitrate* ratio) for each of the test sequences, and each value of source and target *QP*.

In the authors' opinion, the methodology used in this research correctly shows the results of video transcoding. This methodology has also been recently used by the experts of ISO/IEC MPEG in their new standardization activity on network distributed video transcoding.

V. VIDEO QUALITY IN HOMOGENEOUS HEVC TRANSCODING – RESULTS AND DISCUSSION

A. EXPERIMENTAL RESULTS – GENERAL DATA

As evidenced by the results obtained during our experiments, transcoding HEVC bitstreams leads to a noticeable degradation of the quality of video. Figure 3 presents exemplary but representative data of video transcoding that were obtained for the *BlueSky* test sequence. In this figure, the solid line represents the "PSNR – bitrate" curve which is a result of the first step of the experiment: direct encoding of an original test sequence with various *QP*s. A few selected values of $QP = 22, 28, 32, 38$ have been directly marked on the plot using rectangles. This line represents the quality of video C (from Fig. 1) encoded at different *TargetBitrates*.

The dotted lines give information about the results of transcoding of the source bitstream to new bitstreams with different rates. Each dotted line represents the "PSNR-bitrate" curve of the transcoded video produced based on a single selected source bitstream encoded at some *SourceBitrate*. Those source bitstreams were the bitstreams obtained in the first step of encoding for the *QP* values of 22, 28, 32, 38.

The reasons why the quality of the transcoded video is always worse than the quality of the source video, have already been outlined in Section III.

It must be emphasized here that although the overall nature of changes of video quality follow the trends seen for the exemplary sequence (see Fig. 3), the exact results of transcoding also depend on the content of the video to some extent.

The general relationship of quality differences ΔPSNR (quality loss) in the function of the transcoding ratio - *TargetBitrate*/*SourceBitrate* is presented in Fig. 4. In this figure, each dot represents the result of a single transcoding, from a bitstream encoded at the *SourceBitrate* to a bitstream encoded at the *TargetBitrate*. The solid line is the result of averaging the results that are marked with dots. As it can be easily seen, the value of the *TargetBitrate*/*SourceBitrate* ratio affects the results to a large extent.

The main conclusion that can be drawn here is that the lower the value of the transcoding ratio (which translates into a bigger difference between *TargetBitrate* and

SourceBitrate values), the lower the loss of quality expressed by ΔPSNR, as compared to a direct encoding of original data at the *TargetBitrate*.

In the extreme case, in which the *TargetBitrate* is very low, while the *SourceBitrate* is significantly higher (so, the value of the ratio is low as well), transcoding will not lead to a significant degradation of video quality (understood as ΔPSNR). This can be explained by the fact that encoding video material at a very low *TargetBitrate* causes large quality losses. In such conditions, it is not relevant whether we use original video material or one previously slightly degraded by encoding at a relatively high *SourceBitrate*.

In the opposite case of high values of the *TargetBitrate* (which also means higher values of the transcoding ratio), the quality of the video material used (original or previously encoded at a rate similar to the *TargetBitrate*, *SourceBitrate*) makes a significant difference. This highlights a greater loss of video quality during transcoding with a high transcoding ratio, when compared to the scenario with a low transcoding ratio.

Summarizing, we can say that an average maximum PSNR decrease of 1.4dB was observed (in comparison to video encoded directly from original material, at the same *TargetBitrate*) for the range of transcoding ratios considered in the research. One has to remember, though, that this average maximum value also includes results for cases where the target *QP* was lower than the source *QP*, so the *TargetBitrate* was higher than the *SourceBitrate* – a scenario which rarely finds any practical justification. In cases which have the greatest practical use (that is $TargetBitrate/SourceBitrate < 100\%$), the average maximum decrease of PSNR is below 0.7 dB.

Another observation concerns the results of a moderate reduction of the bitrate in the transcoder. What is very interesting, there is an optimal point of transcoding, i.e. an optimal value of the transcoding ratio for which the loss of video quality finds its local minimum in this transcoding scenario. The averaged results (see Fig. 4) show that this point is $TargetBitrate/SourceBitrate = 95\%$ and the loss of video quality is about 0.35 dB for this point. A deeper analysis of this phenomenon will be the topic of subsection V-C.

Apart from the point that ''minimizes'' the loss of video quality (for a moderate reduction of the bitrate), there is also a point of transcoding which forms a local maximum of the loss of video quality. This is for the case of the transcoding ratio approximately equal to 75%, which gives the loss of video quality ΔPSNR = 0.63 dB. Consequently, a reduction of the bitrate by 25% (more or less) results in the highest loss of video quality, when compared to the reference direct encoding results.

Another obvious observation that arises from the analysis of the graph is that the ratios above 100%, for which the *TargetBitrate* is higher than the *SourceBitrate*, are, on average, characterized by a steadily increasing quality loss. The higher the ratio, the higher the loss. The explanation here is that for a higher *TargetBitrate*, transcoding is not able to compete with direct encoding, since direct encoding of original data is able to preserve the details that are simply not presented in the encoded data (video *R* from Fig. 1) fed to the transcoding chain. Therefore, being unable to recover those details, the transcoder is not able to produce a video of similar quality as in the reference scenario.

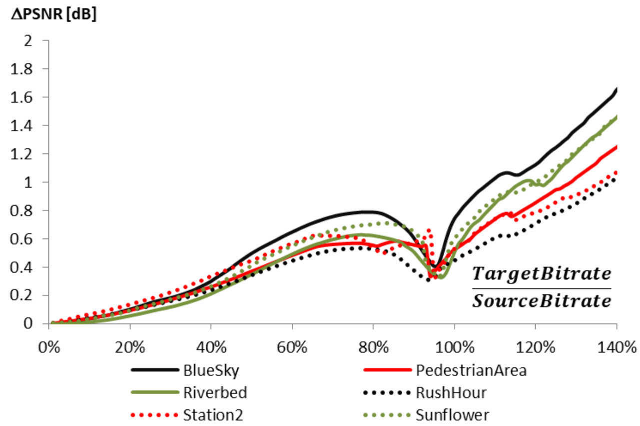

FIGURE 5. The results of averaging the partial data obtained for the considered values of *QP*. Results of individual test sequences.

B. EXPERIMENTAL RESULTS - MORE DETAILED DATA

The exact relationship between the expected ΔPSNR and the transcoding ratio depends on the content of a sequence to some extent but, in general, the overall properties of data remain unchanged. The results that were obtained for individual test sequences are presented in Fig. 5.

A deeper analysis of the presented relationships allowed us to conclude that higher losses of quality occur in more complex (in terms of image content) video sequences. In this case, it is more difficult to predict the content of image blocks which results in higher values of prediction error of image samples that undergo quantization. In the described scenario, the quantization of a prediction error signal causes a larger loss of video quality compared to sequences with a simple content. In the latter case, a smaller prediction error signal is obtained, and that results in smaller losses of information observed in the process of data quantization. It is also worth noting that for complex sequences, even small *QP* (high *SourceBitrate*) used for the first compression in the process, removes a significant amount of data, which are impossible to restore before the second encoding in the cascade transcoder.

The relationship between expected quality loss ΔPSNR and the transcoding ratio depends even stronger on the value of the *SourceBitrate* (i.e., on the value of *QP* used in the first stage of cascade transcoder) than on the content of a video itself. As it can be seen from exemplary results presented for the *BlueSky* sequence (see Fig. 6), the lower the *SourceBitrate* (the lowest bitrate for encoding with $QP = 38$ – round dots, the highest for $QP = 22$ – short dense dashes), the higher the expected losses of video quality when transcoding a video. This is particularly visible for higher values of the transcoding ratio. A more detailed analysis of this property revealed

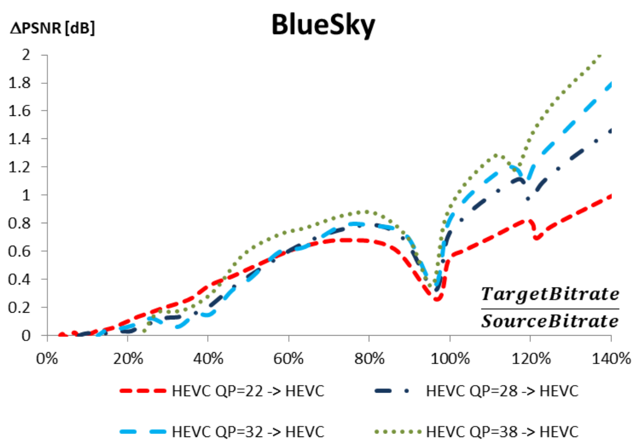

FIGURE 6. **ΔPSNR with respect to transcoding ratio for four scenarios of encoding the original video: $QP = 22, 28, 32, 38$.**

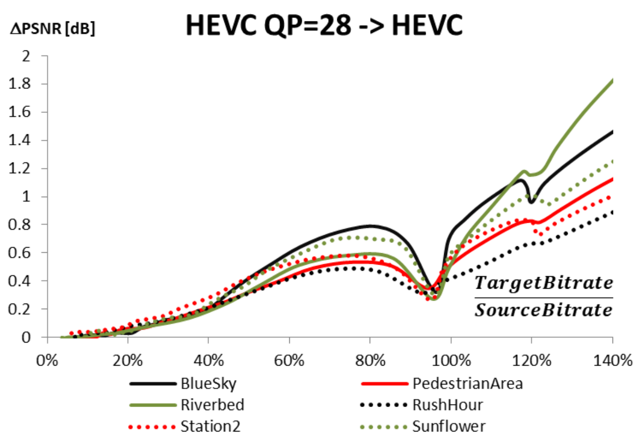

FIGURE 7. **ΔPSNR with respect to transcoding ratio for individual sequences. Results for first step compression with $QP = 28$.**

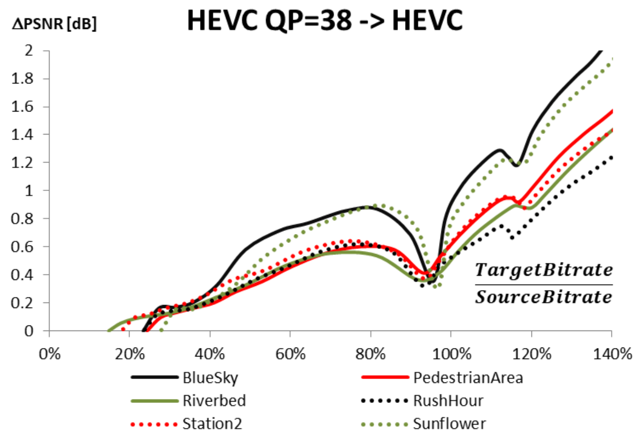

FIGURE 8. **ΔPSNR with respect to transcoding ratio for individual sequences. Results for first step compression with $QP = 38$.**

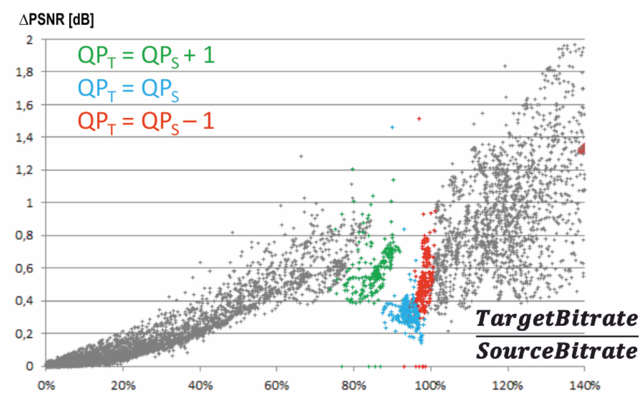

FIGURE 9. **ΔPSNR with respect to transcoding ratio (i.e. *TargetBitrate*/*SourceBitrate* ratio). Figure presents partial data obtained for individual test sequences. With colors results for specific relations of QP_T and QP_S were additionally highlighted.**

that underlying reasons can be found in the nature of the "PSNR-bitrate" curve (see solid line in Fig. 3). This curve is very steep for low values of bitrates (which corresponds to higher values of QP), and relatively flat for high bitrates (so, for low QP values). As a result, transcoding of low bitrate bitstreams will result in a greater change (loss) of video quality, than in the case of higher values of bitrates. Complementary to the results of Fig. 6 are data presented in Figs. 7 and 8, presenting the results obtained for individual test sequences for two strongly different values of QP: 28 and 38. The presented results confirm all the conclusions that were formulated before.

C. CONDITIONS FOR MINIMUM LOSS VIDEO TRANSCODING

As it was pointed out in section V-A, there exists a characteristic point of video transcoding which minimizes the loss of video quality for a scenario of a moderate reduction of bitrate in a video transcoder. According to the averaged experimental data, this point is $TargetBitrate/SourceBitrate = 95\%$. In order to investigate the causes of the existence of such a point, each portion of the partial data has been carefully analyzed in terms of settings of the encoder, which led to individual results. In particular, the relationship between the value of quantization parameter QP_S used for encoding the original video (that produces video R from Fig. 1) and the value of this parameter QP_T used later in the transcoder, was studied. The results of this analysis achieved for the range of the transcoding ratio from 75% to 100% were presented in Fig. 9.

As it turned out, the local minimum in quality loss is obtained for the case when encoding in the second step of transcoding uses the same QP value as the one used in the first step of encoding. The points closest to this minimum correspond to the cases when the second-step encoding was performed with the QP value higher or lower by 1 than the QP used in the first step. Partial results obtained for the various scenarios (mutual relationship of values of QP_T and QP_S) have been marked with appropriate distinctive colors in Fig. 9.

The important thing to note is that the bitrate ratio of 100% is obtained almost exclusively for the cases when encoding in the second step is performed with QP lower by one than the

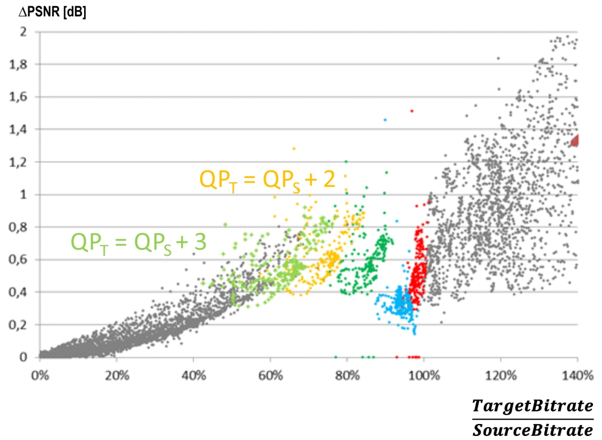

FIGURE 10. **ΔPSNR with respect to transcoding ratio (i.e. *TargetBitrate*/*SourceBitrate* ratio). With green and yellow colors results for specific relations of QP_T and QP_S were additionally highlighted.**

value used in the first step, while causing a loss of 0.5 dB in quality on average.

D. CONDITIONS FOR MAXIMUM LOSS VIDEO TRANSCODING

In the scenario of video transcoding when the ratio *TargetBitrate*/*SourceBitrate* is in the range of 50% to 100% one can easily see cases when quality losses are particularly high. As it can be observed in figures of this section, the maximum losses of video quality were reported for the *TargetBitrate*/*SourceBitrate* ratio approximately equal to 75%. Although it is very difficult to accurately point out the reasons of this phenomenon, we present our thoughts on this matter.

First of all, we noted that the results tend to cluster with respect to the difference between QP_S and QP_T. The maximal loss of transcoding is observed mainly for $QP_T = QP_S + 2$ and $QP_T = QP_S + 3$. This is an important conclusion of the research. The partial results obtained for these cases are shown in the Fig. 10.

Our thoughts on the origin of the presence of the maximum of quality loss is the following.

We observed that when the *TargetBitrate* varies between 80% to 120% of the *SourceBitrate*, the farther the *TargetBitrate* from the point of the minimal losses is, the greater the transcoding loss becomes. This seems to be true regardless of whether the *TargetBitrate* increases or decreases. The mechanism of this increase of loss stems from the misalignment of the quantization steps and does not seem to be surprising. Unfortunately, for *TargetBitrate* lower than 80% of the *SourceBitrate*, we observe an extreme of the quality loss in terms of ΔPSNR.

We understand that this is the result of the fact that the ΔPSNR metric is a superposition of two important factors. One of these factors is the size and overlap of the quantization steps. This mechanism is also used in the transcoder to decrease the bitrate. The other, however, is of a different nature.

For low *Target* to *Source* bitrate ratio, either the *TargetBitrate* itself needs to be low in general, or the *SourceBitrate* needs to be high.

For the first case, the quality of the reference bitstream (coded directly from the original sequence to the *TargetBitrate*) is low, and the difference between the quality of transcoded and reference bitstreams is low (the quality will not decrease much lower than the reference).

For the second case, for which the *TargetBitrate*/*SourceBitrate* is low due to a very high *SourceBitrate*, the points are close to the horizontal axis (i.e., the transcoding loss is low) because the source quality is very high and close to the quality of the original sequence. Then, the transcoding results are very much like the reference coding results, and the transcoding process is not much different (hence, it does not introduce significant loss) than the encoding of the original sequence.

Both of these phenomena cause the transcoding curve to drop (i.e., the transcoding loss is decreasing) after the *TargetBitrate*/*SourceBitrate* decrease below approximately 75%. To sum up, the following seems to be true for the points with bitrate ratio below 75%:

1) The *TargetBitrate* (numerator) is low. Then, both the reference and transcoded sequence are of similar, low quality and ΔPSNR is low.
2) The *SourceBitrate* is high. Then the source stream is of high quality, similar to original sequence. Transcoding in such a case produces similar quality to the quality of the reference encoded bitstream, therefore ΔPSNR is, again, low.

We did not investigate the reason why all these phenomena superimpose to reach the maximum loss at approximately 75% of bitrate ratio, but one important observation is to avoid transcoding with $QP_T = QP_S + 2$ and $QP_T = QP_S + 3$.

VI. VIDEO QUALITY IN HOMOGENEOUS HEVC TRANSCODING - FURTHER CONSIDERATIONS

The results presented in the previous section clearly show that when transcoding a video, the mutual relationship of the quantization parameters of source bitstream QP_S and target bitstream QP_T determines directly how big the loss of quality of a signal will be. Figure 9 presents the empirical data of video quality loss for cases when $QP_T = QP_S$ and $QP_T = QP_S \pm 1$ in a homogeneous HEVC transcoder. But at this point there are still open questions. First, what is the justification of the nature of obtained results? And secondly, how big would the loss of video quality be if the mutual relationship of values of QP_S and QP_T was different from the one already analyzed in section V-C? These issues will be the subject of a more detailed discussion in the following paragraphs.

A. THE NATURE OF OBTAINED RESULTS – ANALYSIS OF REASONS

In the context of the first question, a theoretical analysis has been performed to explore the impact of data quantization on the size of quantization error. This analysis has been carried out assuming a special case of two consecutive stages of quantization. In the considered scenario, the first stage of data quantization is realized with parameter QP_S, while the second stage with parameter QP_T. The following special cases were considered in the analysis:

- Case A: $QP_T = QP_S$,
- Case B: $QP_T < QP_S$,
- Case C: $QP_T > QP_S$.

Conclusions from the analysis have been presented in illustrations of Fig. 11, from which the following general conclusions can be drawn.

The use of two identical quantizers in both stages (thus, $QP_T = QP_S$) will lead to a value of the final quantization error 'e' which will never be greater than the maximum error 'e_{max}' for the second quantizer (see Case A in the Fig. 11). This is because of the equal quantization step sizes used in both the first and second quantizer. Of course, this will be the case when the output value of the first quantizer is directly put onto the second quantizer (such a scenario was adopted in Fig. 11). However, in a real transcoder, the situation is somewhat different. The data that are quantized in a transcoder (with quantization parameter value QP_T) may be different from those quantized in the first stage of video encoding (that is realized with parameter QP_S), due to other steps of processing, like nonlinear filtering and prediction that will influence the final value of the sample. This change of value explains the non-zero loss of video quality for the $QP_S = QP_T$ scenario (as presented in Fig. 9). But the general theoretical conclusion highlighted above that the full compliance of the two quantization step sizes ($QP_S = QP_T$) gives the minimum error of video transcoding is still true, compared with other relationships of values of QP_S and QP_T.

In the case when the two quantizers are different from each other (i.e., QP_S differs from QP_T), the final quantization error 'e' can be bigger than the maximum error 'e_{max}' of the second quantizer. The discussed situation for the two separate cases in which $QP_S < QP_T$ or $QP_S > QP_T$ has been illustrated with Cases B and C in Fig. 11. The presented conclusion results from a different allocation of decision boundaries (on the axis of quantized values) in the first and second quantizer, which comes directly from different widths of quantization intervals. Thus, in these cases the resulting error of transcoding is generally greater than in the case when $QP_S = QP_T$.

B. RELATIONSHIP OF QP_T and QP_S AND ITS IMPACT ON ERRORS OF TRANSCODING

The illustrations from the previous subsection clearly present the mechanism of introducing transcoding errors in a video transcoder when QP_S differs from QP_T, but there is still the unsolved question about the magnitude of this error in practical situations. In order to answer the question of error magnitude for certain cases, the range of values of QP_S and QP_T were tested. This way, the influence of the width difference

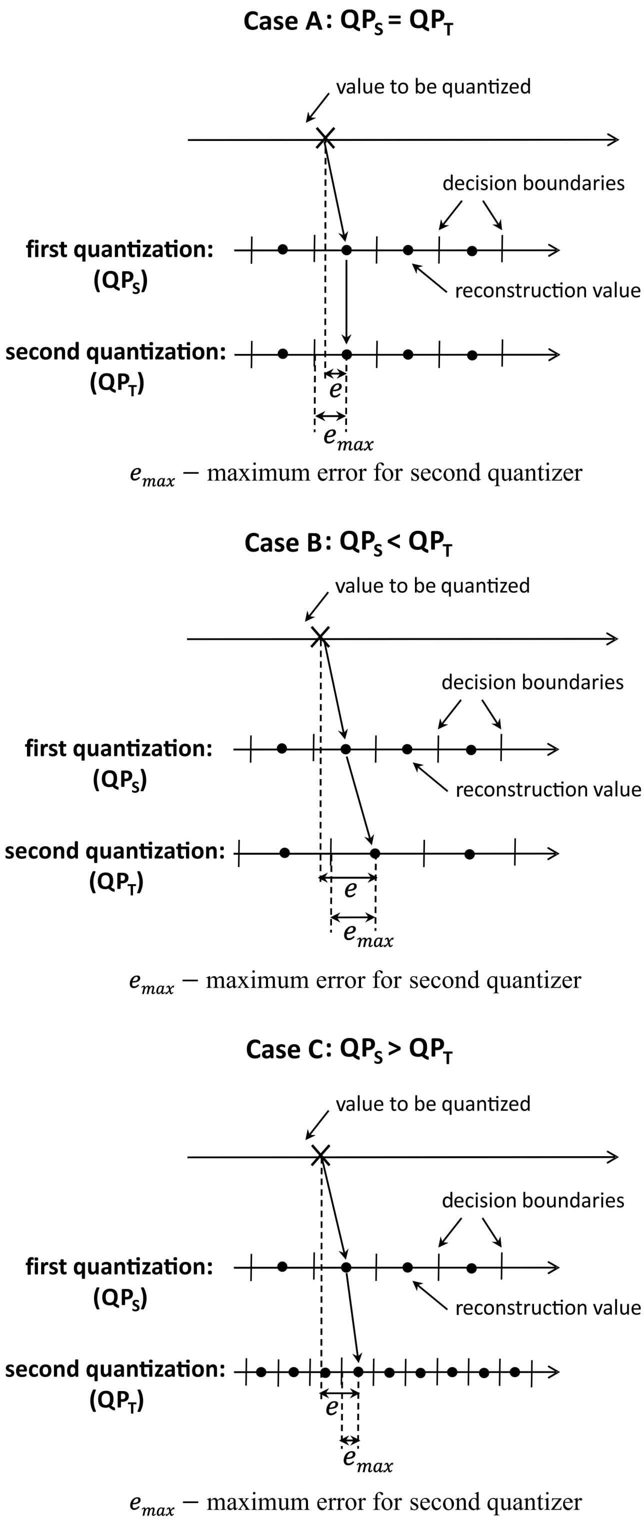

FIGURE 11. Impact of the relation of values of QP_S and QP_T on the size of quantization error. Analysis of three cases: Case A – $QP_S = QP_T$, Case B – $QP_S < QP_T$, Case C – $QP_S > QP_T$.

of the quantization step at different stages of encoding was tested in an additional experiment.

In this experiment, a uniform quantizer with a quantization step equal to $QStep_S$ was used to quantize every possible transform coefficient value from an HEVC encoder.

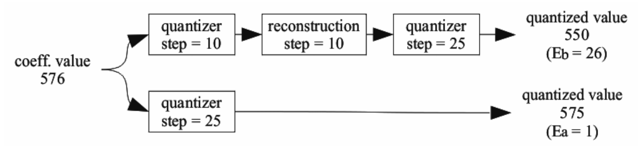

FIGURE 12. An example result of the experiment with varying quantization steps.

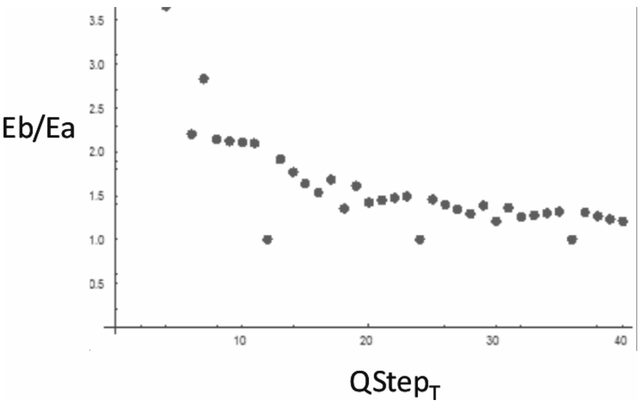

FIGURE 13. The error ratio as a function of $QStep_T$. $QStep_S$ was kept constant at the value of 12.

Then, the dequantized value was quantized once again using a quantization step equal to $QStep_T$. For reference purposes, the same initial value of the transform coefficient was directly quantized using the quantization step equal to $QStep_T$. For both cases, quantization error E was calculated – E_a and E_b respectively for those two cases. The example calculation is shown in Fig. 12.

The average values of errors E_a and E_b for the example case ($QStep_S = 10$ and $QStep_T = 20$) (averaged over all possible coefficient values) are 12 and 14.5, respectively. Therefore, it can be seen that transcoding introduces an additional error when compared to a case with direct compression with a given width of the quantization step, which agrees with the observations described in the previous chapters. The ratio of mean error values (that is Eb/Ea) is 1.2.

To examine the influence of the quantized step width change on the results, a different experiment was conducted, for which $QStep_S$ was kept constant at the value of 12, while $QStep_T$ was changed from 2 to 40 with a step of 1. The ratio of mean error value Eb/Ea is given as a function of $QStep_T$ in Fig. 13.

What is evident from this figure, is the following. The minimum error can be obtained only when the $QStep_T$ width is an integer multiple of the $QStep_S$ width. For such cases the error remains the same in both scenarios, therefore $Eb/Ea = 1$. For every other case, the scenario with requantization gives a higher average error, even when $QStep_T$ changes by 1 from the multiple of $QStep_S$. For such a case, the error can be over twice as high.

Another observation also confirms the previously presented results – the higher the $QStep_T$, the lower the Eb/Ea ratio, and therefore the requantization loss is smaller.

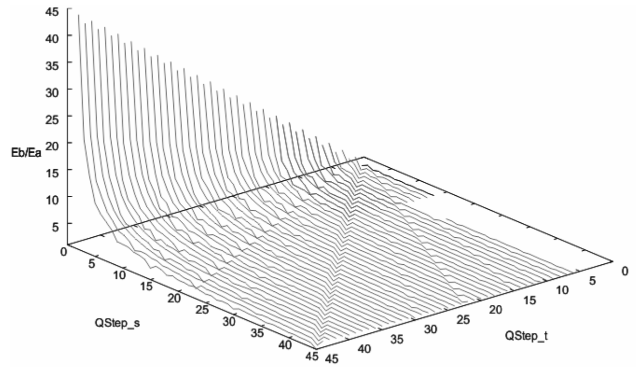

FIGURE 14. The error ratio as a function of $QStep_T$ and $QStep_S$.

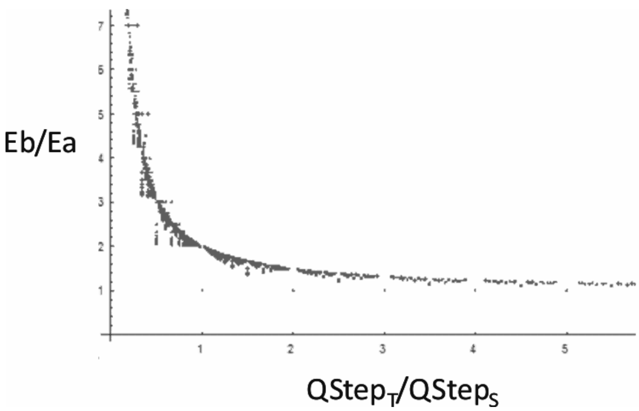

FIGURE 15. The error ratio as a function of $QStep_T/QStep_S$ ratio.

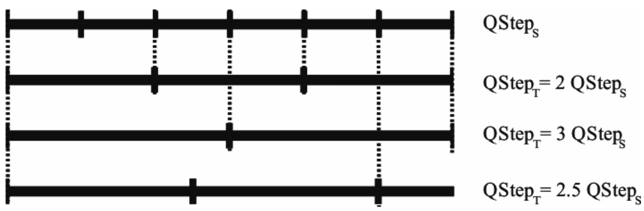

FIGURE 16. The quantization step limits overlap for different $QStep_T/QStep_S$ ratios.

In Fig. 14, the relationship between the Eb/Ea coefficient and both quantization steps, $QStep_S$ and $QStep_T$, is visualized.

The minimum values for $QStep_S = QStep_T$ are clearly noticeable in this figure, as well as a less noticeable minimum for $QStep_T/QStep_S = 1.5, 2.5, 3.5$ and so on.

In Fig. 15, the same data are plotted in 2D – again, the minimums mentioned above are clearly visible.

The reason for the minimum to occur for $QStep_T$ being a multiple of $QStep_S$ is obvious: the quantization step limits (decision boundaries) are overlapping for those cases, therefore the samples are always represented by the same quantized value, regardless of whether they underwent a single quantization with $QStep_T$ or two consecutive quantizations – first with $QStep_S$ and then with $QStep_T$. This is presented in Fig. 16.

The reason for the Eb/Ea ratio for $QStep_T = 2.5 \cdot QStep_S$ not being one is that only every other of the decision

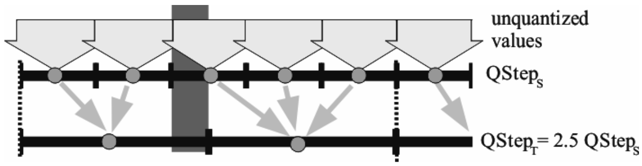

FIGURE 17. Requantization introduced additional error – marked in dark gray.

boundaries overlap. This means that two out of three steps will overlap entirely, but every third one will get requantized and will be assigned to a certain step of $QStep_T$, with a maximal quantization error equal to $0.5 \cdot QStep_S$, as it is shown in Fig. 17. The dark gray area marks the range of input values that are assigned to a different step in the case of re-quantization.

VII. FINAL REMARKS AND CONCLUSIONS

A quantitative analysis of the expected quality losses in a homogeneous HEVC transcoder was presented in the paper. The obtained results can be used as a reference for future research on video transcoder improvements. The results have shown that, although video transcoding always leads to some loss of video quality, there are transcoding conditions for which the loss of quality is minimal - with respect to the quality of video encoded from the original video at the same bitrate as the transcoded video. The minimum quality losses of ΔPSNR = 0.35 dB are achieved with a moderate reduction of the bitrate, for the transcoding ratio $TargetBitrate/SourceBitrate = 95\%$. It corresponds to the case of using the same value of QP in the first and second video encodings. On the other hand, transcoding scenario in which $TargetBitrate/SourceBitrate = 75\%$ gives maximum loses of video quality. This scenario corresponds to a $QP_T = QP_S + 2$ configuration in most cases. However, it should be further noted that the exact relationship of ΔPSNR versus transcoding ratio depends significantly on the value of the $SourceBitrate$, as well as the content of a video sequence.

The results presented in this work, apart from their scientific value, are also of high practical importance. In common industrial practice, content providers prepare multiple versions of the same video material encoded at different bitrates. Currently, the most widely spread-out approach is to prepare video material at halved bitrates (e.g., 2Mbps, 1Mbps, 500kbps, 250kbps), and for small bitrate adjustments of those bitrates video transcoding is utilized. Our analysis clearly shows that such an approach is not always optimal, as the biggest quality losses may occur in the case of such small bitrate adjustments. Instead, based on the presented results, the optimal approach would be to prepare several closely distributed (in bitrates) video streams directly from original material, and to perform video transcoding in a strictly defined scenario of video encoding. The cases when it is worth to use video transcoding (instead of using multiple versions of the encoded video) were indicated in the Fig. 18.

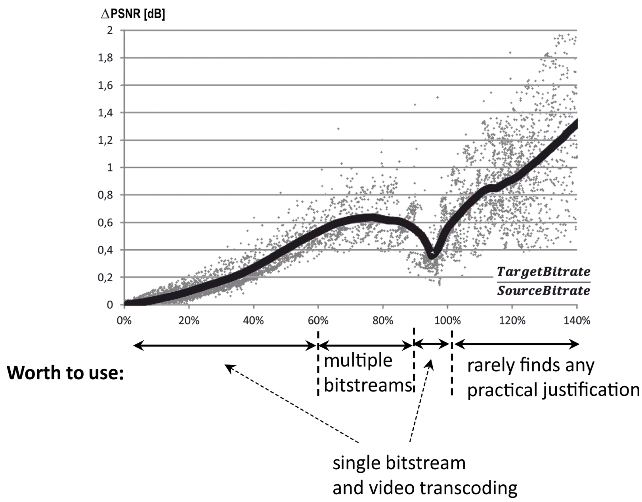

FIGURE 18. Ranges of values *TargetBitrate/SourceBitrate* when it is worth to use video transcoding instead of the solution with multiple bitstreams.

Due to the wide use of cascaded pixel domain video transcoders (CPDT) in industrial practice this technique of transcoding has been considered in the work. The subject of the study was the classical CPDT transcoder that performs the full encoding of a video material. According to the best knowledge of the authors, the obtained results will apply to all of the transcoders that are based on the idea of CPDT, including the state-of-the-art ones. As we suppose, every other type of CPDT transcoder will have the same limitations (those discussed in this work) and possibly, some additional ones, imposed by the method proposed by the authors. Thus, any CPDT transcoder will exhibit the same behavior as described in this paper.

REFERENCES

[1] *Generic Coding of Moving Pictures and Associated Audio Information—Part 2: Video. (MPEG-2)*, document ISO/IEC 13818-2 and ITU-T Rec. H.262, Nov. 1994.
[2] *Video Coding for Low Bit Rate Communication*, document ITU-T Rec. H.263, Aug. 2005.
[3] *Generic Coding of Audio-Visual Objects, Part 10: Advanced Video Coding*, document ISO/IEC 14496-10, Mar. 2006.
[4] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.
[5] *High Efficiency Video Coding—MPEG-H Part 2*, document ISO/IEC IS 23008-2:2015, ITU-T Rec. H.265 Edition 3.0, 2015.
[6] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.
[7] A. Puri, X. Chen, and A. Luthra, "Video coding using the H.264/MPEG-4 AVC compression standard," *Signal Process., Image Commun.*, vol. 19, no. 9, pp. 793–849, Oct. 2004.
[8] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards—Including high efficiency video coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669–1684, Dec. 2012.
[9] T. Grajek, J. Stankowski, K. Wegner, and M. Domański, "Video quality in AVC homogenous transcoding," in *Proc. Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Dubrovnik, Croatia, May 2014, pp. 211–214.
[10] O. Alay and H. Stephansen, "The effect of multi-generation encoding in broadcast contribution on the end-user video quality," in *Proc. 19th Int. Packet Video Workshop (PV)*, Munich, Germany, May 2012, pp. 113–118.

[11] J. Stankowski, T. Grajek, K. Wegner, and M. Domański, "Video quality in multiple HEVC encoding-decoding cycles," in *Proc. 20th Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Bucharest, Romania, Jul. 2013, pp. 75–78.

[12] A. Vetro, C. Christopoulos, and H. Sun, "Video transcoding architectures and techniques: An overview," *IEEE Signal Process. Mag.*, vol. 20, no. 2, pp. 18–29, Mar. 2003.

[13] I. Ahmad, X. Wei, Y. Sun, and Y.-Q. Zhang, "Video transcoding: An overview of various techniques and research issues," *IEEE Trans. Multimedia*, vol. 7, no. 5, pp. 793–804, Oct. 2005.

[14] J. De Cock, S. Notebaert, P. Lambert, and R. Van de Walle, "Requantization transcoding for H.264/AVC video coding," *Signal Process., Image Commun.*, vol. 25, no. 4, pp. 235–254, Apr. 2010.

[15] S. Notebaert, J. De Cock, K. Vermeirsch, P. Lambert, and R. Van de Walle, "Quantizer offset selection for improved requantization transcoding," *Signal Process., Image Commun.*, vol. 26, no. 3, pp. 117–129, Mar. 2011.

[16] S. Moiron, S. Faria, A. Navarro, V. Silva, and P. Assunção, "Video transcoding from H.264/AVC to MPEG-2 with reduced computational complexity," *Signal Process., Image Commun.*, vol. 24, no. 8, pp. 637–650, Sep. 2009.

[17] L. P. Van, J. De Cock, G. Van Wallendael, S. Van Leuven, R. Rodriguez-Sánchez, J. L. Martínez, P. Lambert, and R. Van de Walle, "Fast transrating for high efficiency video coding based on machine learning," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Melbourne, VIC, Australia, Sep. 2013, pp. 1573–1577.

[18] L. P. Van, J. De Cock, A. J. Diaz-Honrubia, G. Van Wallendael, S. Van Leuven, and R. Van de Walle, "Fast motion estimation for closed-loop HEVC transrating," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Paris, France, Oct. 2014, pp. 2492–2496.

[19] L. P. Van, J. De Praeter, G. Van Wallendael, S. Van Leuven, J. De Cock, and R. Van de Walle, "Efficient bit rate transcoding for high efficiency video coding," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 364–378, Mar. 2016.

[20] V. A. Nguyen and M. N. Do, "Efficient coding unit size selection for HEVC downsizing transcoding," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Lisbon, Portugal, May 2015, pp. 1286–1289. doi: 10.1109/ISCAS.2015.7168876.

[21] L. P. Van, J. De Praeter, G. Van Wallendael, J. De Cock, and R. Van de Walle, "Performance analysis of machine learning for arbitrary downsizing of pre-encoded HEVC video," *IEEE Trans. Consum. Electron.*, vol. 61, no. 4, pp. 507–515, Nov. 2015. doi: 10.1109/TCE.2015.7389806.

[22] L. P. Van, J. De Praeter, G. Van Wallendael, J. De Cock, and R. Van de Walle, "Machine learning for arbitrary downsizing of pre-encoded video in HEVC," in *Proc. IEEE Int. Conf. Consumer Electron. (ICCE)*, Las Vegas, NV, USA, Jan. 2015, pp. 406–407. doi: 10.1109/ICCE.2015.7066464.

[23] M. Sung, M. Kim, M. Kim, and W. W. Ro, "Accelerating HEVC transcoder by exploiting decoded quadtree," in *Proc. 18th IEEE Int. Symp. Consumer Electron. (ISCE)*, JeJu Island, South Korea, Jun. 2014, pp. 1–2. doi: 10.1109/ISCE.2014.6884329.

[24] Z. Lin, Q. Zhang, K. Chen, J. Sun, and Z. Guo, "Efficient arbitrary ratio downscale transcoding for HEVC," in *Proc. 30th Vis. Commun. Image Process. (VCIP)*, Chengdu, China, Nov. 2016, pp. 1–4. doi: 10.1109/VCIP.2016.7805465.

[25] W. Kuang, Y.-L. Chan, S.-H. Tsang, and W.-C. Siu, "Fast HEVC to SCC transcoder by early CU partitioning termination and decision tree-based flexible mode decision for intra-frame coding," *IEEE Access*, vol. 7, pp. 8773–8788, 2019. doi: 10.1109/ACCESS.2018.2890720.

[26] C.-M. Fu, E. Alshina, A. Alshin, Y.-W. Huang, C.-Y. Chen, C.-Y. Tsai, C.-W. Hsu, S.-M. Lei, J.-H. Park, and W.-J. Han, "Sample adaptive offset in the HEVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1755–1764, Dec. 2012.

[27] K. Yang, S. Wan, Y. Gong, H. R. Wu, and Y. Feng, "Perceptual based SAO rate-distortion optimization method with a simplified JND model for H.265/HEVC," *Signal Process., Image Commun.*, vol. 31, pp. 10–24, Feb. 2015.

[28] HM 15.0 Reference Software. *Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11*. Accessed: Oct. 2018. [Online]. Available: https://hevc.hhi.fraunhofer.de/svn

[29] F. Bossen, *Common Test Conditions and Software Reference Configurations*, document JCTVC-J1100, Joint Collaborative Team on Video Coding (JCT-VC) of ITUT SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Stockholm, Sweden, Jul. 2012.