Analysis of CABAC context models efficiency in the HEVC

J. Stankowski, D. Karwowski, K. Wegner, O. Stankiewicz , K. Klimaszewski, T. Grajek

[1] Poznan University of Technology, Chair of Multimedia Telecommunications and Microelectronics, POLAND

{jstankowski, dkarwow}@multimedia.edu.pl

Abstract **– This paper presents detailed analysis of the CABAC efficiency when operating in the HEVC in a number of bits per number of bins manner. The paper reveals what fraction of bit is produced by the CABAC for each and every binary symbols encoded in HEVC. The statistics of bins encoded with a given efficiency are presented for the three types of video data: 1) prediction error of image samples, 2) motion data, and 3) control data.**

Keywords – ***CABAC entropy coding, HEVC, Context models***

I. INTRODUCTION

In the final stage of video encoding entropy coding is always used. Entropy coding plays an important role in video compression, because it reduces the statistical redundancy that remains within syntax elements that are produced by video encoder.

In both the AVC and the HEVC video encoders, the syntax elements that are the input for the entropy coding represent various types of video data. In general, these are the following types: 1) prediction error of image samples, 2) motion data, and 3) control data. The three types of data listed above have different statistical properties. In addition, the statistics of the data varies with the change of image content.

The state-of-the-art in the field of entropy coding of video data is Context-based Adaptive Binary Arithmetic Coding (CABAC) [1,2] that has been introduced firstly in Advanced Video Coding AVC MPEG 4 part 10/H.264 standard of ISO/IEC and ITU-T [3], and later in the new High Efficiency Video Coding HEVC MPEG-H part 2 / H.265 international standard of the same organizations [4]. Research on efficiency of CABAC clearly indicate that it ranks among the most efficient entropy encoders used in video compression techniques [1].

As the CABAC is in fact binary arithmetic coder it is able to encode only two binary symbols: 0 and 1. Because of that each syntax element (that are not binary in nature, likes i.e. motion vector differences) needs to be translated into some string of binary symbols. A single binary symbol of this string is called a bin.

In CABAC estimation of probability of each bin is done by the so-called context model. Probability of appearance of 0 and 1 varies and in general depends on actual number of the encoded bin, type of the encoded syntax element and values of the previously encoded bins within a given syntax element.

Therefore, for a given image block, the probability of a bin is estimated with respect to the context information, that holds the information about the values of the given bin in neighboring blocks of the image. Taking into account the possible values of the context information (for a given types of bins), and different types of syntax elements that are defined in HEVC, the whole stream of bins is divided into the total number of 154 sub-streams of bins, in order to track the statistics of data independently in each of the sub-stream [2,4]. Therefore it is also said that CABAC uses 154 independent probability estimators (context models), as illustrated in Fig. 1.

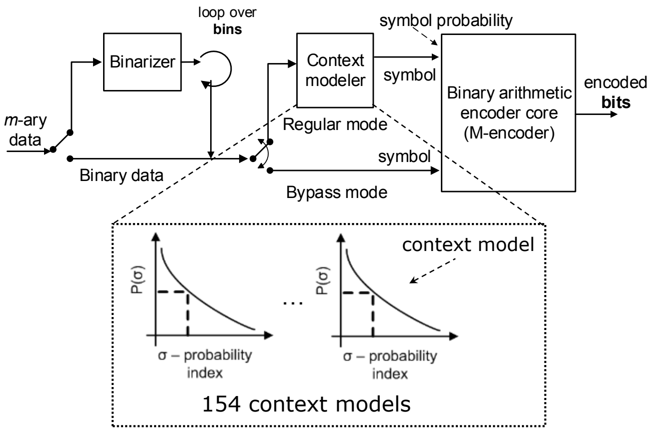

Figure 1. General block diagram of CABAC entropy encoder in HEVC.

All bins are fed into a binary arithmetic encoder core (the so-called M-encoder) which produces bits representing encoded bitstream. It is commonly known that not every bin feed into a binary arithmetic encoder produces a bit at the output. On average CABAC has an ability to represent each bin merely on a fraction of bit.

II. RESEARCH PROBLEM

The goal of this work is to analyze in details for every bin (every 0 and 1 from a stream of bins) what fraction of bit was produced by the CABAC in response to a given bin at the input. This fraction of bit produced in response to a given bin can be understand as compression performance of that particular bin.

The questions we are trying to answer is: How many bins are compressed effectively? How many bins are compressed ineffectively leading to data expansion? What portion of bins is compressed with particular efficiency (i.e. leading to production of 0.2 or 0.3 bit at the output of arithmetic encoder core).

Because we have analyzed each encoded bin separately, conducted experiments gives a large set of experimental data. For this reason, the partial results have been grouped together, first in 154 CABAC context models, and then according to a type of syntax element that have produced a given bin. Following three groups of syntax elements have been used: symbols related to prediction error of image samples, symbols related to motion data and symbols related to control data.

III. EFFICIENCY OF SYNTAX ELEMENT CABAC COMPRESSION – METHODOLOGY

Experiments were done with the HM 10.0 reference software of HEVC [5] configured in the Common Test Conditions (CTC) scenario [6]. In particular, random access test settings were used with a strict division of encoded video into I, and B frame types. In contrast to the CTC guidelines, a wider range of QP values was considered in the research (QP=1-40) when encoding a set of 1920x1080 test video sequences (*bluesky, pedestrianarea, riverbed, rushhour, station2, sunflower, tractor*). The results of the encoding is a comprehensive collection of 280 encoded data streams.

The encoded data streams were then analyzed using the authors' Bitstream Analyzer software. The analyzer performs full CABAC decoding of encoded data streams and calculates the fraction of bit produced by the CABAC in response to each and every bin. The fraction of bit was calculated by the use of the CABAC counter estimator [7]. Gathered fraction of bit produced for every bin was the basis for evaluating performance of CABAC for each of the sub-stream.

Based on the analysis of CABAC efficiency for each of the context models, the bins were divided into separable groups containing bins for which the ratio (defined as: fraction of bit produced by each bin) is contained in a fixed range of values. In particular, the range 0–2 of the ratio was studied, which has been divided into smaller sub-ranges of length 0.2.

IV. EFFICIENCY OF CABAC CONTEXT MODELS – RESULTS

A. *Results for all the CABAC bins*

Among all the coding bins, the largest group makes these symbols, for which the ratio is in the range 0.8–1.0. This group represents 30%–55% of all data symbols. A much smaller proportion of symbols was observed in the other groups. And so, for the 15%–20% of symbols, there was obtained the ratio from range 0.6–0.8, 8%–13% of symbols is encoded with the ratio 0.4–0.6, the ratio in the range 0.2–0.4 was obtained for 5%–12% of the bins, and even lower ratio (0.0–0.2) was obtained for 3%–20% of the bins. In the case of less than 5% of the bins encoder led to the expansion of data, which means that those particular bins was encoded on more than one bit i.e 1.5 bit. Research has shown that the exact proportions of the individual groups of bins in the total data stream depends on the value of QP. With the increase of QP the share of bins for which greater coding efficiency was achieved (i.e. the ratio from range 0–0.8) also increases, and thus the share of bins coded with lower efficiency (the ratio from range 0.8–1) decreases. Detailed results are shown in Fig. 2.

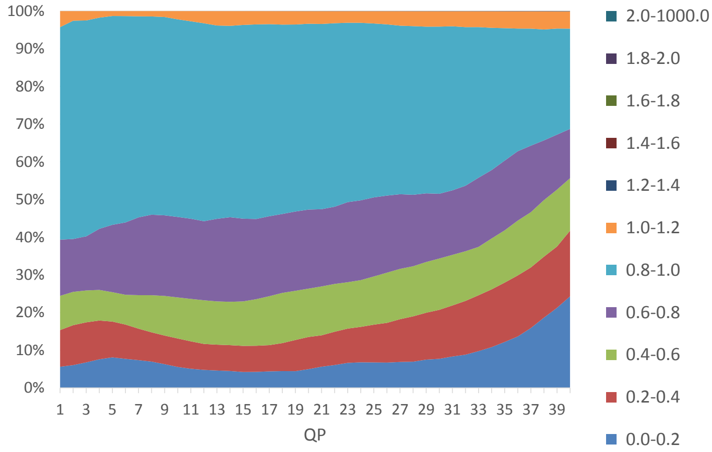

Figure 2. The relative proportion (in %) of the number of bins that have been encoded with the ratio belonging to a given range (0-0.2, 0.2-0.4, etc.). All of the CABAC bins have been considered in statistics.

In order to answer the question why the value of QP affects the results, a detailed analysis of results obtained for individual groups of data is needed. Next subsections present analysis for three main groups of data: 1) data that represent prediction error of image samples, 2) motion data, and 3) control data.

B. *Results for control and motion data*

In the case of both the motion and the control data a strong correlation of proportions of the respective groups of bins with the value of QP was observed. And so, with the increase of QP the share of bins that were encoded with the ratio from range 0–0.8 increases rapidly, thus the proportion of bins for which lower efficiency was achieved (i.e. the ratio from range 0.8–1.0) decreases. In the cases of the control data, 40%–80% of bins are encoded with the ratio 0–0.8, and for 15%–50% of bins the ratio 0.8–1 was achieved. In the case of data representing motion information, the figures are 20%–80% for bins encoded with the ratio 0–0.8 and 10%–75% for bins encoded with the ratio 0.8–1.0. Detailed proportions of the different groups of bins depends on QP and can be read from Fig. 3 and 4. Previous studies of the authors showed that in the case of high QP the pool of coding modes which are actually used by the encoder is strongly reduced [8]. Because the quantization removes the details of the image, the encoder resigns from the use of small image blocks and uses more frequently the blocks of a larger size. Motion estimation performed in larger image blocks results in a more consistent motion field, whereby the prediction of motion data is more efficient (that is, prediction error of motion data often assumes a small value). As a result, the probabilities of data symbols (representing both the control and the motion data) are more varied at higher values of QP. With the increase of QP the entropy of data decreases, which results in higher efficiency of entropy coding.

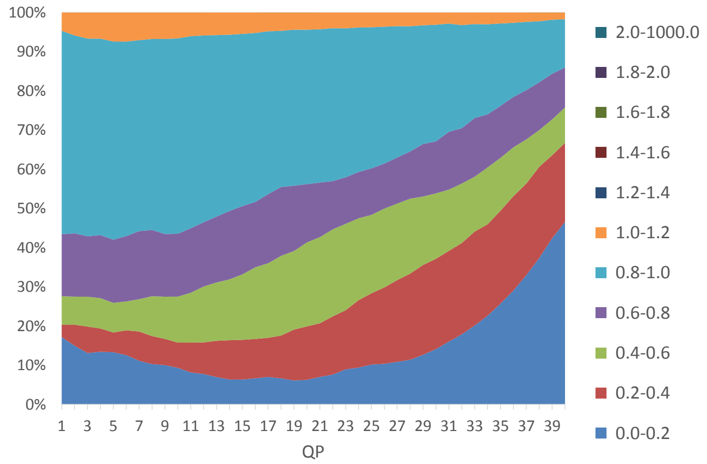

Figure 3. The relative proportion (in %) of the number of bins that have been encoded with the ratio belonging to a given range (0-0.2, 0.2-0.4, etc.). Only CABAC bins that are related to control data were considered in statistics.

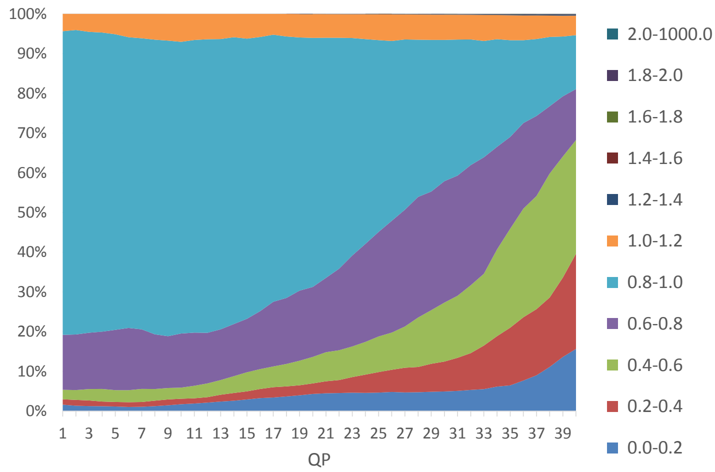

Figure 4. The relative proportion (in %) of the number of bins that have been encoded with the ratio belonging to a given range (0-0.2, 0.2-0.4, etc.). Only CABAC bins that are related to motion data were considered in statistics.

C. *Results for data representing prediction error of image samples*

In contrast to the results presented in subsections A and B, the results for prediction error of image samples do not show strong variation depending on the QP. On average, for half of the data from this group the ratio from range 0–0.8 is achieved. The remainder of the symbols are coded with the ratio 0.8–1.0 (about 40% of symbols), and symbols for which the number of bits for each and every bin encoded is greater than 1 i.e 1.5 bits per bin (less than 10% of such symbols for which the entropy coding is carried to the expansion of data). Detailed results were presented in Fig. 5.

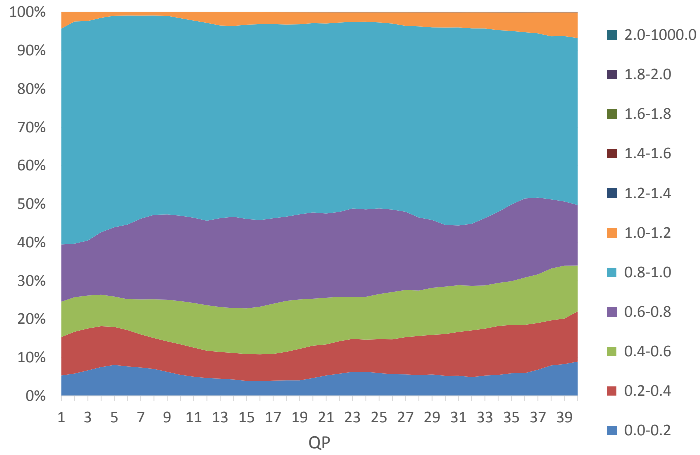

Figure 5. The relative proportion (in %) of the number of bins that have been encoded with the ratio belonging to a given range (0-0.2, 0.2-0.4, etc.). Only CABAC bins that are related to information about prediction error of image samples were considered in statistics.

Prediction of image samples is in the HEVC encoder performed in blocks of a certain size. In general, sizes of blocks from 4x4 to 64x64 are available. The used in the encoder selection criterion of a block size is to minimize the cost of a block encoding, which generally refers to minimization of prediction error of block samples. Thus, for each value of QP the encoder selects the coding mode that minimizes the entropy of image sample prediction error. For this reason, similar efficiency of entropy coding was observed for this group of data, regardless of the QP.

V. CONCLUSSIONS

The paper presents research on the efficiency of the CABAC entropy encoder in HEVC. For each of the 154 context models, performance has been measured as an fraction of a bit produced by each and every encoded bin. It has been shown that such data give detailed information about the efficiency of the CABAC entropy encoder in HEVC. For the sake of presentation, the data has been aggregated and analyzed in groups, which document the efficiency of the algorithm for the three types of video data. The results show that for the vast majority of the bins CABAC performs efficient compression. Although the detailed results depend on the type of encoded data and value of QP (as depicted in section IV), the averaged results show, that for half of the bins the ratio from the range of 0–0.8 bit was achieved, and for most of the other bins the ratio was 0.8–1 bit. What is interesting, there are also some bins, which encoding leads to expansion of the data. In the extreme case, these bins constituted 5% of all the bins. This gives a strong indication that some of CABAC context models can be coded more efficiently.

The results attained during the research may be a basis for future works on development of the CABAC improvements as well as fast mode decision algorithms in a video encoder.

A. *Fast caclulation of cost of application of a coding mode in a video encoder*

For each of image block the encoder selects the compression mode, which is a compromise between the number of bits that are sent to the decoder and the quality of reconstructed block. Thus, for each of the tested modes the abovementioned parameters must be calculated. From the amount of computations point of view determining the bit cost of a mode is more difficult. Accurate determination of the number of bits requires realization of full encoding path, including the entropy coding of data. Realization of all the CABAC functional blocks at the stage of coding modes testing gives accurate results, but such a solution is associated with very long encoding time.

For the purpose of calculating the bit cost for a mode, there is a possibility of significant simplification of the CABAC encoder, without losing on the accuracy of calculation of the cost. With the attained raw results for fraction of bit produced for each and every bin (averaged for each of the 154 CABAC context models) one can determine the bit cost of a mode with high accuracy, relying only on the data observed at the output of the binarizer block. So, we completely ignore the parts of statistical modeling and the core of arithmetic coding when testing the successive compression modes. The available studies on the complexity of the CABAC encoder show that the proposed solution will significantly shorten the encoding time of a sequence [9]. The authors are currently conducting research in the field of the proposed solution.

B. *Improvement of the CABAC entropy encoder*

The raw research results for fraction of bit per bin allow an accurate assessment of the efficiency of each of the 154 CABAC context models. This knowledge can be the basis for further improvements of the algorithm oriented to increase the efficiency of entropy coding. In this context, the improvement of those syntax elements is crucial, for which the fraction of bit per bin exceeds 0.8. In this group of bins, there are also those for which there is expansion (and not compression) of the data (see Fig. of section IV).

Presented in Fig. 3 averaged results for the control data highlights, that there can be even 60% of the bins that are encoded with the ratio of 0.8 or higher. In the case of the motion data, the same situation may take place for even 80% of the bins (see Fig. 4). The things look better for the prediction error of image samples data, for which there are about 50% of the bins that are encoded with the ratio 0.8 or higher. This indicates for great importance of the works on more efficiently encoding of the motion and control data.

ACKNOWLEDGMENT

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

REFERENCES

[1] D. Marpe, H. Schwarz, and T. Wiegand: Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 7, pp. 620-636, July 2003.

[2] V. Sze, M. Budagavi, "High Throughput CABAC Entropy Coding in HEVC," IEEE Transactions on Circuits and Systems for Video Technology, vol.22, no.12, pp. 1778-1791, Dec. 2012.

[3] ISO/IEC 14496-10, Generic Coding of Audio-Visual Objects, Part 10: Advanced Video Coding, March 2006.

[4] ISO/IEC 23008-2 (MPEG-H Part 2) / ITU-T Rec. H.265: High Efficiency Video Coding (HEVC), Apr. 2013.

[5] HEVC test model reference software – available online: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/

[6] F. Bossen, Common test conditions and software reference configurations, Joint Collaborative Team on Video Coding (JCT-VC) of ITUT SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Doc. JCTVC-J1100, Stockholm, Sweden, July 2012.

[7] F. Bossen, CE1: Table-based bit estimation for CABAC, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Document: JCTVC-G763, Geneva, Switzerland, November 2011.

[8] J. Stankowski, T. Grajek, D. Karwowski, K. Klimaszewski, O. Stankiewicz, K. Wegner, M. Domański, Analysis of Frame Partitioning in HEVC, Lecture Notes in Computer Science – Computer Vision and Graphics, pp. 602-609, Springer International Publishing Switzerland 2014.

[9] D. Karwowski, Computational complexity analysis of adaptive arithmetic coding in HEVC video compression standard, Image Processing and Communications, September 11-13, 2013, Bydgoszcz, Poland.