Bitrate Distribution of Syntax Elements in the HEVC Encoded Video

Jakub Stankowski, Damian Karwowski, Tomasz Grajek, Krzysztof Wegner, Jakub Siast, Krzysztof Klimaszewski, Olgierd Stankiewicz, Marek Domański

Chair of Multimedia Telecommunications and Microelectronics
Poznan University of Technology, POLAND
{jstankowski, dkarwow, tgajek, kwegner, jsiast}@multimedia.edu.pl

***Abstract*— The paper focuses on two problems: 1) analysis of bitrate contribution of I and B frames in encoded HEVC data stream, and 2) analysis of distribution of HEVC syntax elements in I and B frames. The relationship between sizes of I and B frames located at a different levels in hierarchical coding scheme was experimentally determined, as well as contribution of HEVC syntax elements in encoded streams of I, B0, B1, B2, and B3 frames. The results are presented in the paper. All presented results concern a wide range of QP values in a video encoder.**

***Keywords*— HEVC; video compression; video data analysis; video bitstream content**

I. INTRODUCTION

Currently, more than half of the total traffic in telecommunication networks is related to video. Without any doubts, this factor will steadily increase in the next years. This raises the need for an efficient compression of a video in order to reduce the data transmission costs. Among the known techniques of a video compression, of particular importance are hybrid coding schemes of a video with intra-frame and inter-frame prediction followed by lossy transform coding of residual image data [1]. The hybrid coding scheme has been included in many international and commercial recommendations (MPEG-2, H.263, AVC, VC-1, AVS, HEVC), and find now widespread applications.

The state-of-the-art in the field of hybrid video coding is the new High Efficiency Video Coding technology defined in international recommendation of ISO/IEC and ITU-T [2]. The HEVC technology is a breakthrough in a video compression and enables significant reduction of a bitstream representing encoded images [3]. Although the main idea of encoding as well as the general structure of the video encoder remained unchanged when compared to previous solutions (e.g. MPEG-2, AVC), there are a lot of improvements and new coding tools that distinguish the HEVC from its predecessors [4]. As a result, the encoded data stream of HEVC is different and not compatible with data streams that are produced by older encoders.

Significant differences in the stream structure pose a problem for efficient HEVC encoder control, since the older solutions, for older encoders (e.g. MPEG-2, H.263, MPEG-4 AVC) do not work as expected. These solutions are extremely specialized, since the base of their action is the knowledge of the structure of encoded data that is produced by encoder e.g. [5],[6]. Compared to the predecessors, different syntax elements are produced in the HEVC encoder. Quantitative features of data that form a bitstream are also different in the case of HEVC. As a result, bitrate control methods that are used in older encoders cannot directly be used in the new HEVC encoder. Experiments on the structure of a video bitstream, already done for older encoders, must be repeated for the HEVC technology. Thus, the topic of our research was to analyze the HEVC bitstream properties and contents.

II. RESEARCH PROBLEM AND GOAL OF THE WORK

This work is a contribution to further study on efficient bitrate control in the HEVC encoder. The goal of the work is to perform a thorough analysis of data stream that is produced by the HEVC encoder. Results obtained in this field up to date and presented in the literature do not give answers for many detailed questions. For example, in [7] results obtained for I and B frames were averaged. There are no separate results for individual frame types. Additionally, experiments were conducted only for two QP values. Moreover above mentioned results are gathered using early version of HM software and do not correspond to final HEVC standard. In [8] authors focused on the average number of context/bypass coded bins without detailed analysis for separate syntax elements.

Our goal is to examine the following issues:

- the relationship between the size of images of a different types in encoded data stream (I- and B-frames),
- the relationship between the size of individual images that belong to different encoding levels in hierarchical coding of images i.e. I-, B0-, B1-, B2-, B3-frames (see explanation in the next section) and use of B frames,
- the distribution of syntax elements in encoded stream of data representing I-, B0-, B1-, B2-, B3-frames,
- influence of the value of quantization parameter (QP) on the abovementioned properties.

III. Methodology of the Research

This work focuses on analysis of HEVC encoded data stream that will probably be a typical stream to be used in the future home theatre systems and broadcasting environments. Therefore, the specific assumptions (presented below) about the methodology of experiments were made.

Research works were done using the commonly available HM 10.0 reference software of the HEVC standard [9]. Using this software, a series of experiments was done with the following encoding conditions.

- Seven 1920x1080 test video sequences were used: *bluesky*, *pedestrianarea*, *riverbed*, *rushhour*, *station2*, *sunflower*, *tractor*. The sequences are recommended by ISO/IEC MPEG as well as ITU-T VCEG working groups as appropriate for the purposes of research works on video compression.
- 'Random access' encoding scenario was used, as defined in 'common test conditions' [10]. In particular, a video sequence was divided into Groups Of Pictures (GOP), with a strict division of a GOP into I-, B-frame types and hierarchical encoding of B-frame types (the so-called B0, B1, B2 and B3 frames) within GOP (see Fig. 1. for details).
- Experiments were done for a wide range of target bitrates by changing the value of quantization parameter (QP) in encoder. Tests were done for QP from 10 to 40 with step equal to 1.
- Other settings of the encoder were set as suggested in 'common test conditions'.

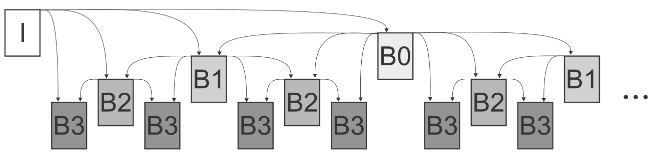

Fig. 1. Hierarchical coding scheme. Only part of prediction sources are marked.

As a result of research, 320 encoded data streams were achieved. Each of data streams was the subject of careful examination, which was conducted for the purposes of:

- Analysis of contribution of individual frames (I, B0, B1, B2, B3) in the encoded data stream, as a function of QP value.
- Detailed analysis of contribution of HEVC syntax elements in individual frames (I, B0, B1, B2, B3), depending on the value of QP.

The partial results were then averaged in order to draw more universal conclusions. The volume of data that represents the partial results does not allow for their inclusion in this paper. Only averaged results were presented in the paper.

The partial results have been published on the companion website to this paper: *www.multimedia.edu.pl/icses2014-hevc*.

IV. Experimental Results

In the first phase of experiments, the general information about distribution of the bits among individual frames of HEVC was collected. The results are presented in Fig. 2 as a function of QP value.

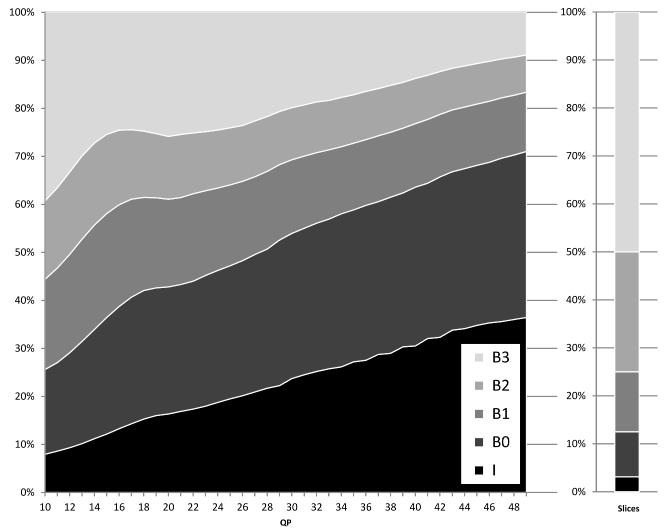

Fig. 2. Contribution of individual frames in encoded data stream, averaged over 7 test sequences results for I, B0, B1, B2, and B3 frames (left). Contribution of frames in total number of slices (right).

In adopted scenario of experiments, the number of I, B0, B1, B2, and B3 frames constituted 3.1%, 9.3%, 12.5%, 25%, 50% of all encoded frames, respectively (it results from the assumed structure of GOP during experiments). Nonetheless, from the viewpoint of the amount of data, the I frames still make a large part of encoded data stream (from 10% to 30% depending on the value of QP). This result is comparable to contribution of B0 frames in a stream – from 15% to 40% depending on the QP value. In this way, I and B0 frames make a significant part of the encoded data stream – from 25% to 70% of the whole data stream depending on the value of QP. The rest of the bitstream is formed by the remaining frames: about 20% for B1 frames, about 15% for B2 frames and 10%-40% for B3 images (depending on QP value).

Presented above knowledge about distribution of data between I and B frames can be only a basis for frame-level mechanism of bitrate control in encoder. Therefore, in the second phase of experiments, the individual frames were additionally thoroughly analyzed in terms of syntax elements distribution in a frame. Detailed statistics obtained for I, B0, B1, B2, and B3 frames were presented in Fig. 3 below for a wide range of QP values.

The main observation is that a type of frame (I or B) as well as the value of QP affect strongly quantitative results.

In the case of I frames, the essential part of the bitstream is formed by three types of data: transform data for luma (70% for lower QPs and 55% for higher QPs), transform data for chroma (20% for lower QPs and 10% for higher QPs), and intra prediction mode (10% -15% depending on value of QP).

In the case of B frames (B0, B1, B2, B3) there is higher number of syntax elements that form a bitstream. Among the syntax elements present in B frames, the following types of data should be distinguished: transform data for luma (80% for lower QPs and 20% for higher QPs), transform data for chroma (20% for lower QPs and 1% for higher QPs), intra and inter prediction modes (2%-7% and 2%-20% respectively), merge index (1% for lower QPs and 40% for higher QPs) and skip flag (1%-9% depending on value of QP). The other types of data make relatively small part of the image data stream.

Besides the value of QP, detailed statistics are obviously dependent on the hierarchy of a B frame in a GOP. What is worth noting, the amount of merge index and skip flag data in B2 and B3 frames is higher relative to that for B0 and B1 frames. At the same time, the amount of transform data is smaller in B2 and B3 when compared to B0 and B1 frames.

The results obtained for QPs in the neighborhood of 30 are of particular importance for high quality video encoding scenario. Bar charts in Fig. 3 present these results in more detailed way. In this case, the essential part of I frame falls on transform data for luma (about 65%), transform data for chroma (about 16%), and information on intra prediction mode (10%). Distribution of data in B0, B1, B2, and B3 frames is relatively aligned: transform data for luma makes 65%-72% of total bitstream while the inter prediction mode signaling makes 7% to 10% of the bitstream. Notable difference between individual B frame types lie in contribution of transform data for chroma as well as skip flag.

For comparison purposes, the influence of hierarchical QP value in the case of B0, B1, B2, and B3 frames was presented in the second bar chart on Fig. 3. The value of QP affects the amount of transform data in the final stream.

The following conclusions can be drawn from the presented results:

- separate mechanisms of bitrate control should be used for I and B frames due to different statistics of syntax elements,
- the statistics for B0 and B1 frames are similar to each other, so one common mechanism of bitrate control for these frames can be used,
- the results for B2 and B3 correspond well, so one method of bitrate control can be used for both, B2 and B3 frames.

V. Conlusions and Final Remarks

Analysis of the HEVC bitstream content was documented in this paper.

The experimental data give useful information about the size of individual frames in encoded stream, but also information about distribution of syntax elements in frames of a given type. Valuable result of the work is a huge collection of partial results which are accessible via companion website: www.multimedia.edu.pl/icses2014-hevc. They are an important contribution to the knowledge about efficient coder control mechanisms for the new HEVC standard.

On the basis of detailed results, two general conclusions can be highlighted. First, I and B0 frames make a significant part of the encoded data stream: 25% to 70% of the total bitstream (exact value depends on QP). Second, four syntax elements constitute an important part of data stream: transform data for luma and chroma, mode of prediction, merge index and skip flag. Detailed values depend on image type and QP value. When compared to AVC, the residual data is significantly larger part of the bitstream (HEVC: 75% on average, AVC: 65% on average). The contribution of motion data is very similar in both standards (~10%). The results presented in the paper may be a contribution to further study on efficient HEVC bitrate control.

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] B. Girod, "The Efficiency of Motion-Compensating Prediction for Hybrid Coding of Video Sequences," Selected Areas in Communications, IEEE Journal on , vol.5, no.7, pp. 1140-1154, Aug 1987.

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

[3] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T.-K. Tan; T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC)," IEEE Transactions on Circuits and Systems for Video Technology, vol.22, no.12, pp. 1669-1684, Dec. 2012

[4] G.J. Sullivan, J.-R. Ohm, W.-J. Han, T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard", IEEE Transactions on Circuits and Systems for Video Technology, vol.22, no.12, pp. 1649-1668, Dec. 2012.

[5] T. Grajek, M. Domańnski, "Single Frame Rate-Quantization Model for MPEG-4 AVC/H.264 Video Encoders", Lecture Notes in Computer Science - Computer Vision and Graphics, pp. 384-391, Springer Berlin Heidelberg, 2010.

[6] T. Grajek, M. Domanski, "New model of MPEG-4 AVC/H.264 video encoders," 17th IEEE International Conference on Image Processing (ICIP), pp. 961-964, Sept. 2010.

[7] J. Lainema, K. Ugur, A. Hallapuro, "Single entropy coder for HEVC with a high throughput binarization mode", Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, Doc. JCTVC-G569, Geneva, Switzerland, Nov. 2011.

[8] 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.

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

[10] F. Bossen, " 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, Stockholm, Sweeden, Jul. 2012.

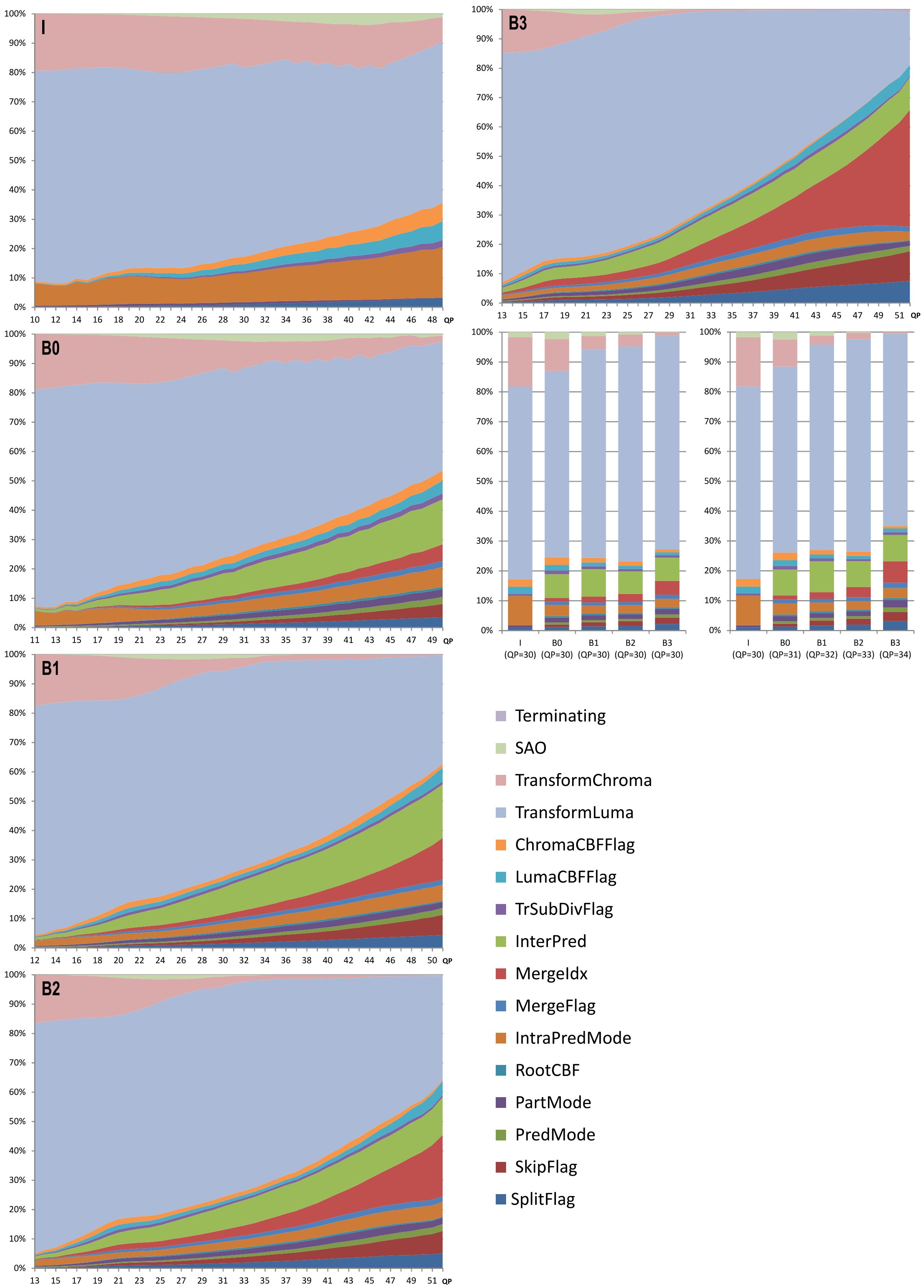

Fig. 3. Contribution of individual HEVC syntax elements in encoded **I**, **B0**, **B1**, **B2**, **B3** frames (averaged results from 7 test sequences).