High Efficiency Video Coding HM-16 for B Frame Comparison: Case Study

ZORAN MILICEVIC (1), ZORAN BOJKOVIC (2)
[1]Department of Telecommunication and IT GS of SAF, 11000 Belgrade
[2]University of Belgrade, Studentski trg 1, 11000 Belgrade
REPUBLIC OF SERBIA
E-mail: [1]zmilicko@yahoo.com, [2]z.bojkovic@yahoo.com

Abstract: - In order to continuing our previous work in the area of performance analysis for High Efficient Video Coding (HEVC) concerning importance parameters such as signal-to-noise ratio (SNR), bit rate and encoding time test models HM-16.20 and HM-16.18 are considered through two test sequences, i.e. Bosphorus (3840x2160) and Jockey (3840x2160). It is shown that there are no differences in SNR and bit rate values. As for the encoding time saving, there are negligible differences. The obtained results are verified through the corresponding simulation. Results for all tested sequences processed by YUV player, as well as Elecard_hevc_analyzer are provided, too.

Key-Words: - Bit-rate, encoding time saving, High Efficient Video Coding (HEVC) standard, HEVC test model (HM-16.20), Main profile, signal-to-noise ratio (SNR).

1 Introduction

It is well known, that High Efficient Video Coding (HEVC) standard is design and developed to focus on two key issues: increased video resolution and increased use of parallel procedure architecture. Description of the HEVC is provided in the open literature [1]-[6]. Although the models of HEVC exist in the H.264/AVC standard, almost everyone has been reconsidered, while many new state-of-the-art coding tools have been introduced. HEVC also presents many new changes including encoding optimization, mode decision, rate-control, hardware design, and error concealment. Next, it should be noted that computational complexity and high memory bandwidth are required in real time processing. On the other hand, the encoding computational complexity of HEVC depends of the test model (HM) configuration. The increased computational complexity leads to high energy consumption and increased power dissipation. Also, the processing time is of huge interest during the coding that requires a video to be transmitted in a limited period as well as is continuous data streaming. This is one of the main reasons for processing time to be in the focus during real time video information transmission.

This paper is organized as follows. After short introduction, background and motivation, the main part of this paper deals with the analysis of the HEVC test models: HM-16.20 vs. HM-16.18. SNR values together with bit rate and encoding time are taken into consideration. Experimental results obtained through simulation processes including the corresponding discussion are in the focus of the second part of this presentation.

2 Short HEVC Background

The video coding layer of HEVC employs the same hybrid approach (inter-/intrapicture prediction and 2-D transform coding) used in all video compression standards since H.261 [3].

The various features involved in hybrid video coding using HEVC are highlighted as follows [3]: Coding tree units and coding tree block (CTB) structure; Coding units (CUs) and coding blocks (CBs); Prediction units and prediction blocks (PBs); Transform units (TUs) and transform blocks; Motion vector signaling: Advanced motion vector prediction (AMVP); Motion compensation; Intrapicture prediction; Quantization control; Entropy coding: CABAC; In-loop deblocking filtering; Sample adaptive offset (SAO).

A number of design aspects new to the HEVC standard improve flexibility for operation over a variety of applications and network environments and improve robustness to data losses. However, the high-level syntax architecture used in the H.264/MPEG-4 AVC standard has generally been retained, including the following features [3]: Parameter set structure, NAL unit syntax structure, slices, Supplemental enhancement information (SEI) and video usability information (VUI) metadata.

Finally, three new features (Tiles, Wavefront parallel processing, Dependent slice segments) are introduced in the HEVC standard to enhance the parallel processing capability or modify the structuring of slice data for packetization purposes. Each of them may have benefits in particular application contexts, and it is generally up to the implementer of an encoder or decoder to determine whether and how to take advantage of these features [3].

Many advanced coding tools have been newly adopted in HEVC to improve the coding efficiency. Among these, HEVC adopted a hierarchical coding structure, which is one of most powerful tools to improve coding efficiency of HEVC. The hierarchical coding structure of HEVC is based on the quad-tree structure of coding unit (CU) where each CU block and has the prediction unit (PU) blocks of symmetric or asymmetric sizes and transform unit (TU) blocks of quad-tree partitions.

For intra prediction, HEVC specifies 35 different prediction modes for luma samples. In HEVC, there are 33 angular modes, a DC mode and an interpolation mode.

Inter prediction, or motion compensation, is conceptually very simple in HEVC, but comes with some overhead compared to H.264/AVC [5]. Similar to H.264/MPEG-4 AVC, HEVC supports quarter sample precision motion vectors. HEVC also supports multiple reference pictures, and the concepts of I, P, and B slices are basically unchanged from H.264/MPEG-4 AVC [5].

The time H.264/AVC was under standardization, a few devices supported high-definition (HD) videos. As a successor of H.264/CAV, HEVC was designed with the goal to satisfy emerging demands of high-quality video services, for example HD TV [3]. The basic processing unit in HEVC is longest coding unit (LCU) which is no overlapped squared block. Each coding unit (CU) can be divided into four partitions.

Today, researchers are exploring the way how to reduce the HEVC encoders complexity. The focus has been in reducing the motion estimation (ME) complexity, because ME occupies 77%-81% of HEVC encoders implementation. Performance comparison of HEVC with older standards such as H.264/AVC, MPEG-4 Part 2 Visual, H.262/MPEG Video, H.263 and also with image coding standards such as JPEG, JPEG 2000, was carried out [7]. Also, researchers are exploring transcoding between HEVC and other standards such as MPEG-2 and H.264. Further extensions of HEVC are scalable video coding (SVC), 3D video multiview video coding and range extensions which include screen content coding (SCC), bit depth larger than 10 bits and color sampling of 4:2:2 and 4:4:4. In general, SCC refers to computer generated object, both images and videos requiring lossless coding. Some of these extensions have been finalized by the end of 2014, while time frame for SCC was late 2016. Iguchi et al. have already developed a hardware encoder for super hi-vision (SHV) i.e., with HDTV at 7680x4320 pixel resolution. Also real time hardware implementation of HEVC encoder for HD video has been done.

The increased encoding complexity represents one of the very important challenges for real-time applications. The quad-tree structure for coding unit with different sizes and a large number of prediction modes is one of the reasons for encoding complexity of HEVC. Thus, one of the challenges for real-time applications is to develop a test mode decision method for reducing computational complexity for HEVC. Secondly, several different methods have been investigated recently, aiming of computational complexity reduction and scaling of HEVC software implementations. Thus, maintaining the encoding time for frame or group of pictures (GOPs) below an adjustable upper bound is still an open research issue.

During the HEVC standardization process, the JCT-VC also developed a reference Software HEVC test model (HM). The aim of the reference software was to provide a basis upon which to conduct experiments in order to determine coding performance. In HM, pictures are first divided into slices, while slices are divided into sequence of treeblocks. A treeblock is a square block (64x64 pixels) of luma samples together with two corresponding blocks of chroma samples. The coding unit (CU) is a basic unit of the splitting region used for inter/intra predictions. The CU concept allows treeblock recursive splitting into four equally sized blocks. This process generates a content-adaptive coding tree structure comprised of CU that may be as large as a tree block as small as 8x8 pixels. The prediction (PU) is the basic unit used for caring the information related to the prediction processes. During the HEVC standardization, the HEVC test model reference software adopted same fast encoding algorithm [8]-[9].

Also, during the HEVC background an analysis the question often arises is: How to improve usability? In order to answer this question, some procedures have to be taken over, i.e

- Support for encoding in terms of latency and quality,
- Improved quality for video coding,
- Support for constant bit rate CBR, variable bit rate VBR and adaptive variable bit rate (AVBR),
- Streamlined quality measurement for 4K 10-bit coded video based on peak signal-to-noise ratio (PSNR),
- Performance measurement on core platforms.

This offering supports a wide range of applications, services, eco-systems and devices [3-5].

The next release understands:

- Improved of subjective quality by content adaptive partitioning and mode decision,
- Improved quantization for HEVC encoding including persistence based quantization,
- Implemented Human range of interest and its support via detection of Face/Skin tones,
- Added Human Visual System sensitive psycho-quantization to be integrated.

Along with adding these tools, there is a need to add a track able basis for quality measurement that may be closer to subjective quality.

3 Motivation

The motivation of this work is to present and highlight the latest developments and analysis of HEVC-related technologies implementations and systems. Also, it is of interest to provide users with a deep understanding of this emerging video coding standard and the related state of-the-art technologies. There are a few questions often arising. First of all is what about the SNR, bit-rate and encoding time saving values when HM-16.20 vs. HM-16.18 standards are used? Secondly, what are the results of subjective video assessment for the corresponding tested sequences processed by YUV player? Of course, the third question is related to analyses of total amount of compression values using Elecard_hevc_analyzer tool. In order to answer these questions the simulation conditions have to be explained before representing the performance.

4 Simulation Results

Simulation results represent the next step of our experimental work on performance evaluation for various versions of HM software test model in different conditions [10].

We evaluated the performance of the HEVC model HM-16.20 [11] vs. HM-16.18 [11], when encoder_lowdelay_main configuration is used. The system platform was the Intel(R) Core(TM) i3-4030U Processor of speed 1,9 GHz, 4 GB RAM, and Microsoft Windows 8.1 Enterprise. The HEVC software configurations were as follows: Main profile, value of Level: 5.0 for Jockey and Bosphorus, I picture, B pictures, period of I-pictures: only first (of both configurations), Hadamard transform was used, Maximum coding unit depth was 4, MV (Motion Vectors) search range was 64, SAO (Sample Adaptive Offset), AMP (Asymmetric Motion Partitions) and RDOQ (Rate-Distortion-Optimized Quantization) were enabled, GOP (Group of Pictures) length 4 in encoder_lowdelay_main for IBBB format was used. The QP (Quantization Parameter) used was 32.

Processed configuration is adopted in the reference test models to Main profile.

Experiments were carried out on the tested sequences with fixed quantization parameter value QP=32. We chose QP=32 as value of the QP, because it is approximately average value in reference software’s setup configuration.

For the experiments two different test sequences are selected. The selected test sequences are in the same resolution and frame rates. We used the first 50 frames of test sequences Jockey and Bosphorus. The both test sequences are in the resolution 3840x2160 pixels. All the test videos are in YUV 4:2:0 format and progressive. Test sequence Jockey represents sequence when horse racing is shown with panning to the left to follow [12]. Test sequence Bosphorus represents sequence where luxury yacht is zoomed, with huge bridge on background and camera panning to the right to follow [12].

Also, the SNR values of luma (Y) component of pictures are used. We measured SNR only for Y because human visual system is more sensitive to luma then to chroma components of pictures.

Three fundamental parameters such as: signal-to-noise ratio, bit rate and time saving will be taken into consideration and compared because they are essential for improving usability in HEVC respectively.

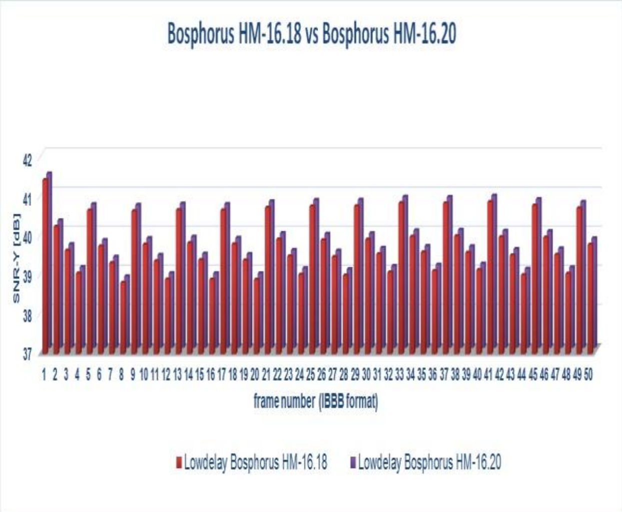
Table 1 shows the performance and comparison of the reference software versions of HM-16 codec for B pictures processing in the IBBB format in lowdelay configuration for QP=32, respectively, based on our simulation results. In the IBBB format 98% of all pictures belong to B picture, while the rest is reserved for I picture.

Table 1. Performance and comparison of the test sequences based on simulation results

Test sequences (resolution)	Profile Main Configuration	Parameters	Reference software HM-16.18	Reference software HM-16.20	Results
Bosphorus (3840x2160)	Lowdelay	SNR-Y (dB)	39,83	39,83	0,00
		Bit-rate (kbps)	1961,79	1961,79	0,00
		Time saving (sec)	20218,94	20250,95	0,16
Jockey (3840x2160)	Lowdelay	SNR-Y (dB)	38,96	38,96	0,00
		Bit-rate (kbps)	1440,94	1440,94	0,00
		Time saving (sec)	20537,39	20483,50	-0,26

When both reference test models of HM-16 codec are tested and compared for Bosphorus test sequence there are not differences in SNR values for luma component of picture, as it is represented in Fig. 1 (a).

a)



b)

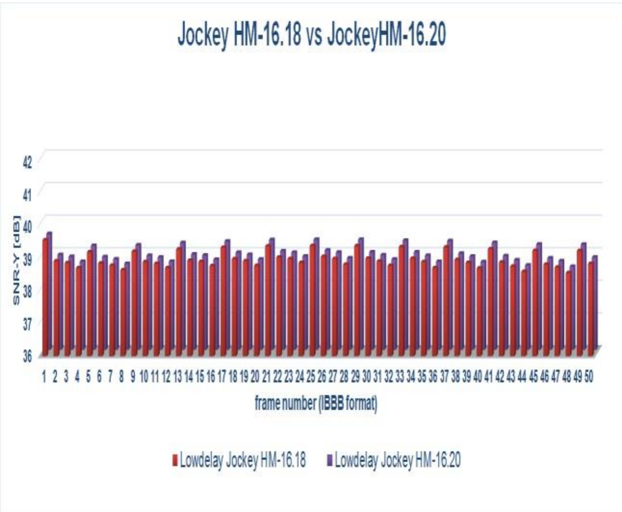


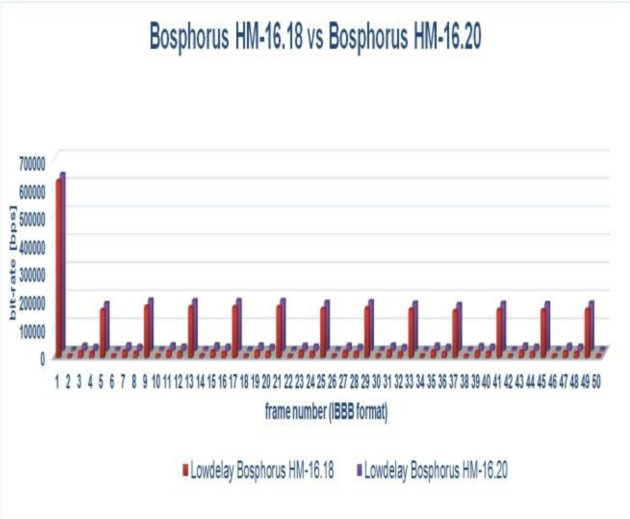
Fig. 1. The SNR curves for Jockey and Bosphorus when lowdelay configuration in IBBB picture format is processed and compared in the both reference test models of HM-16 codec.

The same results are obtained when Jockey test sequence are processed through both reference test models HM-16 codec, as it is shown in Fig. 1 (b).

From the bit rate point of view, for both test sequences when reference test models of HM-16 codec are tested and compared there are not differences in bit rate values for luma component of picture in tested configuration, too.

Fig. 2 (a) represents simulation results for Bosphorus test sequence, while Fig. 2 (b) represents results for Jockey test sequence.

a)



b)

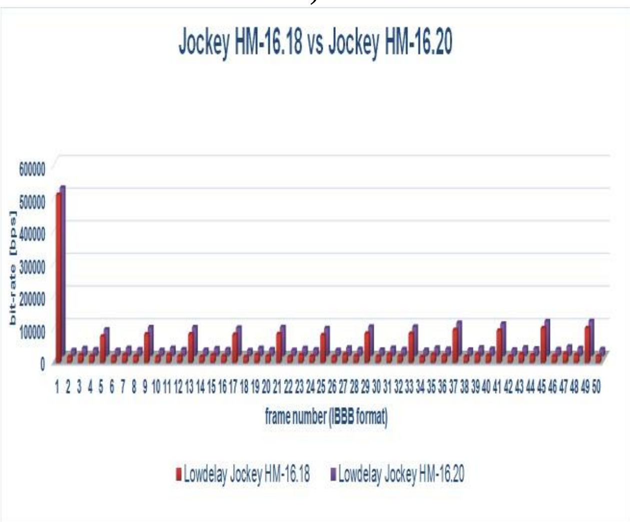


Fig. 2. The Bit-rate curves for Bosphorus and Jockey test sequences when lowdelay configuration in IBBB picture format is processed and compared in the both reference test models of HM-16 codec.

Finally, for Bosphorus test sequence the encoding time saving is increased negligible when both reference test models of HM-16 codec are compared. On the other hand, encoding time saving for Jockey test sequence decrease slightly (denoted by „-”) when HEVC model H.16-20 is compared with HM-16.18.

Next, we used Elecard_hevc_analyzer to analyze total amount encoder compression when the performance of the HEVC reference test models HM-16.18 and HM-16.20 were evaluated for encoder_lowdelay_main configuration.

In Fig. 3 ((a) and (b)) the segment of the decoded video frames are shown, when both test sequences are processed through Elecard_hevc_analyzer for both compared HEVC reference test models of HM-16 codec HM-16.18 and HM-16.20.

a)

BOSPHORUS HM-16.20

BOSPHORUS HM-16.18

b)

JOCKEY HM-16.20

JOCKEY HM-16.18

Fig. 3. The segment of the decoded video frames when both test sequences are processed through Elecard_hevc_analyzer for both compared reference test models of HM-16 codec HM-16.18 and HM-16.20.

When Bosphorus test sequence is analyzed, from total amount of values for intra prediction share is the same 2,07%, while for inter prediction is 12,38% and for transform is 69,48%.

On the other hand, when Jockey test sequence is analyzed, from total amount of values for intra prediction share is the same 4,97%, while for inter prediction is 18,26% and for transform is 54,29%.

Besides objective analysis of the HEVC encoders for two different resolution test sequences, subjective video quality is analyzed, too.

Fig. 4 ((a) and (b)) shows HEVC HM-16.20 vs. HM-16.18, respectively, for both tested sequences, configuration and IBBB picture format for subjective video assessment, respectively. All tested sequences are processed by YUV player, respectively. Subjective assessment results clearly indicate that there are not differences in term of SNR in Fig. 1 in accordance with results in Table 1.

a)

BOSPHORUS HM-16.20

BOSPHORUS HM-16.18

b)

JOCKEY HM-16.20

JOCKEY HM-16.18

Fig. 4. HEVC subjective video assessment for Bosphorus (a) and Jockey (b) test sequence when lowdelay configuration in IBBB picture format is processed in both reference test models of HM-16 codec.

5 Conclusion

The results presented in this paper indicate HEVC standard HM-16.20 vs HM-16.18 when lowdelay configuration is used for testing two test sequences in the same resolution. The SNR, bit-rate and encoding time saving are measured for test sequences in the same resolution when B picture format (IBBB) is processed. Simulations results have shown that there are not differences in SNR values and bit rate values when both reference test models of HM-16 codec are compared for Bosphorus and Jockey test sequences. Also, from the encoding time saving point of view, there are negligible differences reference test models of HM-16 codec. Next, percentage of intra and inter prediction as well as transform in total amount encoder compression values are analyzed, too. Finally, results of subjective video assessment for all tested sequences processed by YUV player are provided, when performance for HEVC HM-16.20 vs HM-16.18 encoder are analysed.

References:

[1] I.E. Richardson, Coding Video: A Practical Guide To HEVC And Beyond, Wiley, 2017.

[2] High Efficient Video Coding, document ITU-T Rec.H-265, April 2013.

[3] G.J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand., Overview of the High Efficient Video Coding (HEVC) Standard, *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 22, No. 12, 2012, pp. 1649-1668.

[4] G. Correa et al. ,Performance and computational complexity assessment of high efficiency video encoders, *IEEE Trans. on Circuits and Systems Video Technology*, Vol. 11, No.12, 2012, pp. 1899-1909.

[5] F. Bossen, B. Bross, K. Suhring,and D. Flynn, HEVC Complexity and Implementation Analysis, *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 22, No. 12, 2012, pp. 1685-1696.

[6] J-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, Comparison of the Coding Efficiency of Video Coding Standard – Including High Efficient Video Coding (HEVC), *IEEE Trans. on Circuits and Systems for Video Technology*, 2012, Vol. 22, No. 12, 2012, pp. 1669-1684.

[7] V. Sze, M. Budagavi and G. J. Sullivan (Editors), High Efficiency Video Coding (HEVC): Algorithms and Architectures, Springer Cham Heidelberg New York Dordrecht London, Springer International Publishing Switzerland 2014.

[8] R.-H. Gweon, Y.-L. Lee, J. Lim, "Early Termination of CU Encoding to Reduce HEVC Complexity", document Joint Collaborative Team on Video Coding of

ISO/IEC, JCTVC-F045, Turin, Italy, July 2011.
[9] J. Jang et al., "Early Skip Detection of HEVC", document Joint Collaborative Team on Video Coding of ISO/IEC and ITU-T, JCTVC-6543, Geneva, Switzerland, November 2011.
[10] Zoran M. Miličević, Zoran S. Bojković, Analysis of High Efficiency Video Coding (HEVC) test model HM-16.18 for P frame processing, *26nd Telecommunications forum TELFOR 2015*, Serbia, Belgrade, November 20-21, 2018, pp. 603-606.
[11] svn://hevc.kw.bbc.co.uk/svn/jctvc-hm/, 2019.
[12] http://ultravideo.cs.tut.fi/#testsequences, 2019.