PERFORMANCE EVALUATION OF H.265/MPEG-HEVC, VP9 AND H.264/MPEG-AVC VIDEO CODING

Ahmad A. Mazhar

College of Computing and Informatics, Saudi Electronic University, Taif Campus, Taif, Kingdom of Saudi Arabia.

ABSTRACT

This study evaluates the performance of the three latest video codecs H.265/MPEG-HEVC, H.264/MPEG-AVC and VP9. The evaluation is based on both subjective and objective quality metrics. The assessment metric Double Stimulus Impairment Scale (DSIS) is used to evaluate the subjective quality of the compressed video sequences. The Peak Signal-to-Noise Ratio (PSNR) metricis used for the objective evaluation. Moreover, this work studies the effect of frame rate and resolution on the encoders' performance. The extensive number of experiments are conducted with similar encoding configurations for the three studied encoders. The evaluation results show that H.265/MPEG-HEVC provides superior bit-rate saving capabilities compared to H.264 and VP9. However, VP9 shows lower encoding time than H.265/MPEG-HEVC but higher encoding time compared to H.264.

KEYWORDS

H.265, HEVC, H.264, VP9, DSIS.

1. INTRODUCTION

Video coding is widely used in a variety of applications such as TV streaming, online gaming, virtual reality tours and video conferencing. These applications require good compression techniques so the communication bitrate is reduced without compromising the quality. In the field of video compression, the H.264 has been dominating on many video applications since it was released in 2003 [1]. The H.264 showed high coding efficiency and reliability especially for standard-definition streaming with limited data rate channels available. The codec with its reference software JM was evaluated and tested in many publications and showed that the H.264 achieved about 50% in coding efficiency compared to its earlier versions [2,3].

The H.264 and VP8 are designed mainly for resolutions lower than High Definition (HD);however, the resolutions nowadays and in the near future demand codecs that are designed to support HD resolutions in addition to Ultra High Definition (UHD) that means 3840x2160 and more.

The huge demand on higher resolutions led to propose a lot of new techniques and compression improvements on video coding. One of the most popular codecs High-Efficiency Video Coding (HEVC) that was released in its first edition in 2013. This codec was a development of ITU-T VCEG and ISO/IEC MPEG Joint Collaborative Team on Video Coding. The proposed HEVC is

DOI : 10.5121/ijma.2016.8103

designed to be valid for all H.264 application with additional supports on high-resolution videos. TheHEVC isofficially approved by ITU-T as H.265 Recommendation and as MPEG-H, Part 2 by ISO/IEC [4].The new H.265-HEVCcodec achievements were significant and substantially higher bit-rate savings with almost same visual quality compared to preceding H.264 codec [5, 6].

Video compression is an open competition area and many codec developers are working apart the team of ITU-T and ISO/IEC. The giant Google company has also an important share in the field of video compression by its codec so-called VP8. Since Google announced about the dedicated "WebM" project in 2010 for high quality open media for the web, the open source video compression technique VP8 is used as the core of the project [7, 8]. The VP8 format was initially developed by a small team On2 Technologies Inc. after it'srecentlyownedby Google Inc. Google started developing VP9 in 2011 as an improved successor of VP8 and released in 2013 [9,10].

VP9 and H.265 are competing and claiming that they provide more efficient compression than H.264. However, VP9 is less popular than H.265 and has little comparative studies to evaluateits efficiency compared to other popular codecs. In the direction of providing a reliable and scientific decision, this paper presents an intensive number of experiments to provide subjective and objective evaluation beside a discussion of each codec of the three existing state of the art.

This paper is organized as follows:Section 2 introduces the selection of codecs implementation. Section 3 discusses the methodology of the study. The experimental results with detailed analysis are discussed in Section 4. Finally, the paper is concluded in Section 5.

2. ENCODER IMPLEMENTATION SELECTIONS

2.1 H.265/HEVC Encoder

The H.265 encoding was evaluated and studied in many publications. The HEVC encoder is considered one of the most popular and efficient encoders [11, 12]. The FFmpegx265 is the implementation used in this study to conduct the evaluation of the codec performance.

2.2 VP9 Encoder

The VP9 encoder was released by Google in both bit stream format and the encoder. The latest encoder version has two-pass run options. The encoder options improved the rate-distortion performance. The FFmpeglibvpx (VP9)is used in this study where the full encoder details are discussed in Section 3.

2.3 H.264/MPEG-AVC

The open encoder x264 is generally used to evaluate the H.264 encoder. In 2006 the first x264 version was released and proven as reliable and efficient and it is currently adoptedin many applications [13]. The x264 provides both command line interface and graphical user interface (GUI). The HandBrake and FFMPEG are the latest interfaces used to implement and run x264 encoder. The latest x264 versions contain some enhancements and features that increase encoding qualitycompared to other H.264 encoders [14]

3. METHODOLOGY OF THE STUDY

3.1 Encoders' Parameters

This section shows the detailed analysis of the encoders used in the study in addition to the evaluation environment. As mentioned in the preceding sections, x265 reference software was used for implementing HEVC encoder where libvpx and x264 were used for implementing VP9 and H.264/MPEG-4 encoder respectively. The selection of encoding settings were determined carefully for fair comparison. In order to evaluate the differences in encoders' capabilities, similar settings were used for all tested encoders.For HEVC, the chosen encoder version is FFmpeg x265 with Main Profile. The full chosen parameters of x265reference software are shown in Table 1. For additional coding parameters, the proven parameters by [15] were also adopted for better coding efficiency and computational complexity reduction.

Table 1. HM Encoder Parameters

R/D optimization	Enabled
Profile	Main
Motion Estimation	EPZS
Group of Pictures (GOP)	8
Deblocking Filter	Enabled
Hadmard Enabled	Enabled
Fast Encoding	Enabled
Bit Depth	8
Fast Merge Decision	Enabled
Rate Control	Disabled
Transform Skip Fast	Enabled
Search Range	64
Reference Frames	4
Intra Period	One second
QP	24, 28, 32, 36

For the x264 encoder, the QP values were set to be 24, 28, 32 and 36. The Group of pictures were also set similar to HEVC. The full setting of x264 is shown in table 2. The settings of libvpxare shown in Table 3 where the 2-pass recommended settings were used as recommended in [16]. The VP9 parameters were set to be similar to HM and x264 parameters. To get fair comparison with variety values, QPs were set to 24, 28, 32 and 36.

Table 2. x264 Parameters

R/D optimization	Enabled
Profile	Main
Motion Estimation	EPZS
Group of Pictures (GOP)	8
Deblocking Filter	Enabled
Hadmard Enabled	Enabled
Fast Encoding	Enabled

Bit Depth	8
Fast Merge Decision	Enabled
Rate Control	Disabled
Transform Skip Fast	Enabled
Search Range	64
Reference Frames	4
Intra Period	One second
QP	24, 28, 32, 36

Table 3. VP9 Parameters

Threads	0
Profile	0
Lag in frames	25
QP	24, 28, 32, 36
Cq-level	20
Auto alt-reference	1
Passes	2
Drop frame	0
Static threshold	0
Undershoot-pct	100
Codec	VP9
Arnr-maxframes	7
Minsection-pct	0
Maxsection-pct	2000

3.2 Testing Environment

For running variety of experimental results, various video sequences were selected with different resolutions and frame rates. The tested sequences were chosen according to the popular higher resolutions and frame rates in current TV broadcasting and Internet streaming. All sequences were tested on computers with Intel core i7 CPU, 4GB RAM and 19'' LED screen. For objective evaluation, only 150 frames were selected from each sequence for simplicity. For the subjective evaluation, ten seconds of each video sequence was shown to the viewers.

3.3 Objective Evaluation Study

The rate-distortion performance (RD) assessment was evaluated using Bjontegaard-Delta bit-rate (BD-BR). The BD-BR measurement is a widely used in in calculating average bit-rate differences between R-D curves for the objective quality. The negative values of BD-BR are indicating on the actual saved bit-rate [17].

3.4 Subjective Evaluation Study

The Double Stimulus Impairment Scale (DSIS), Variant II was selected to implement the subjective quality assessment experiments. In DSIS, the original sequence with the decompressed one are presented to be evaluated by a viewer. The viewer is asked to rate the annoyance in a level from 1-5. The detailed implementation of DSIS measurement is described and recommended by ITU-R BT. 500-11 ITU [18]. For obtaining subjective evaluation results,ten-second test video sequences were shown to thirty viewers, where different spatial andtemporal resolutions were considered. The original sequence was shown to the viewerand then the decompressed video sequence. After processis repeated, the viewer can register the impressionon the compression using a five-level quality scale. All viewers were youth with good communication skills. The experiments were in a computer labusing 19" Dell LCD monitors. The lab's windows were covered with gray curtains with white lightening. All participants used the same lab and monitor specifications. The test was running over three days in equal session length to avoid viewer exhausting which may affect their evaluation ability.

To describe the score of the subjective evaluation, statistical measures were computed for all tested resolutions and frame rates. The Mean Opinion Score (MOS) is selected and calculated for this description.

4. EXPERIMENTAL RESULTS AND ANALYSIS

The comparative study with detailed analysis is presented in this section. Table 4 shows the video sequences that were selected for this performance evaluation study. Different frame resolutions and frame rates were chosen in order to provide wider evaluation study.

Table 4. Test Video Sequences

Sequence	Frame resolution	Frame rate/fps
People on street	2560×1600	40
Traffic	2560×1600	30
Basketball drive	1920×1080	40
Kimono	1920×1080	30
Johnny	1280×720	30
Kristin and Sara	1280×720	24
China speed	1024×768	24

4.1 Objective Experimental Results

The detailed results with the calculated (BD-BR)and Time Saving (TS)are presented in Table 5 and Table 6 respectively. At the same PSNR, the ΔBD-BR and ΔTS were calculated as the average of the four selected QP values (24, 28, 32 and 36). The negative values of BD-BR and TS indicate to savings in bit-rate and coding time respectively. In contrast, the positive values indicate the increment. The average BD-BR savings of the x265 compared to x264 is 37.19%. However, x265 outperformed libvpx 48.39% in average BD-BR savings. As also observed in Table 5, x264 performed around 29.6% better savings in BD-BR compared to libvpx.

Table 5. Bit-rate saving comparison between HEVC, H.264 and VP9

ΔBD-BR (%)			
	x265 vs. x264	x265 vs. libvpx	x264 vs. libvpx
People on street	-25.6	-45.3	-22.3
Traffic	-36.1	-52.4	-26.5
Basketball drive	-41.6	-49.7	-30.2
Kimono	-37.6	-42.4	-21.7
Johnny	-44.5	-55.1	-39.6
Kristin and Sara	-42.8	-50.2	-35.4
China speed	-32.1	-43.6	-31.5
Average	-37.19	-48.39	-29.60

Figure 1 presents the BD-BR curves of the three codecs in relative to each other. As shown in the figure, x265 outperformed x264 and libvpx. The bit-rate saving of x265 relative to libvpx is higher than it relative to x264. The x264 outperformed libvpx and provided higher bit-rate saving in all different resolutions and frame rates.

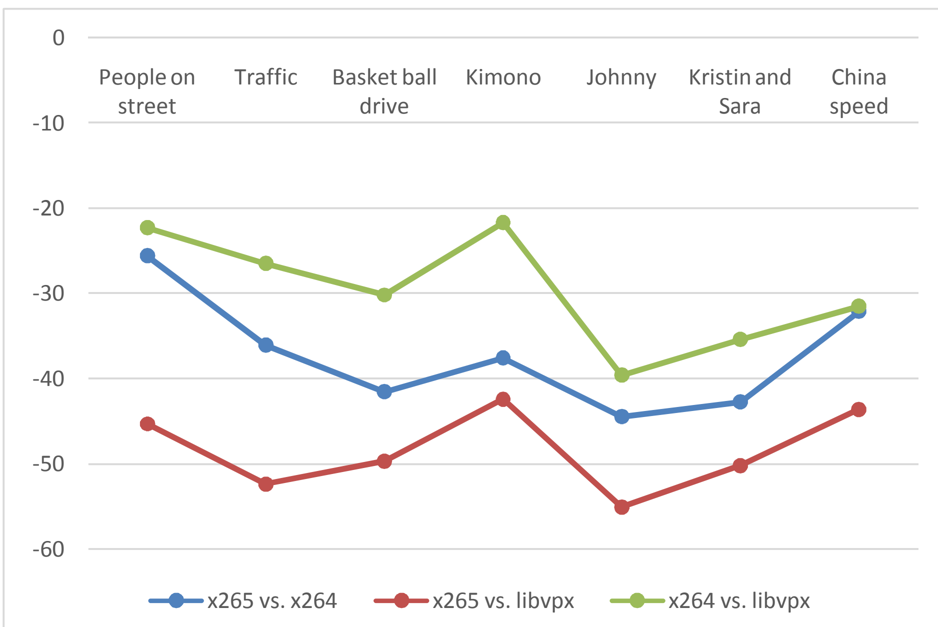

Figure 1. Bit-rate saving comparison between HEVC, H.264 and VP9

In term of saving the encoding time, x264 showed lower encoding time than x265 and libvpx as shown in Table 6. The x265 and libvpx encoding time is higher than x264 about 107.4 times and 140.69 times respectively. However, x265 is lower encoding time than libvpx by a factor of 6.59.

Table 6. Time saving comparison between HEVC, H.264 and VP9

ΔTS (%)			
	x265 vs. x264	x265 vs. libvpx	libvpx vs. x264
People on street	11820	361	7322
Traffic	17315	562	12483
Basketball drive	11290	603	9810
Kimono	13870	672	11021
Johnny	17613	792	13943
Kristin and Sara	12879	801	9782
China speed	13701	826	10827
Average	14069.71	659.57	10741.14

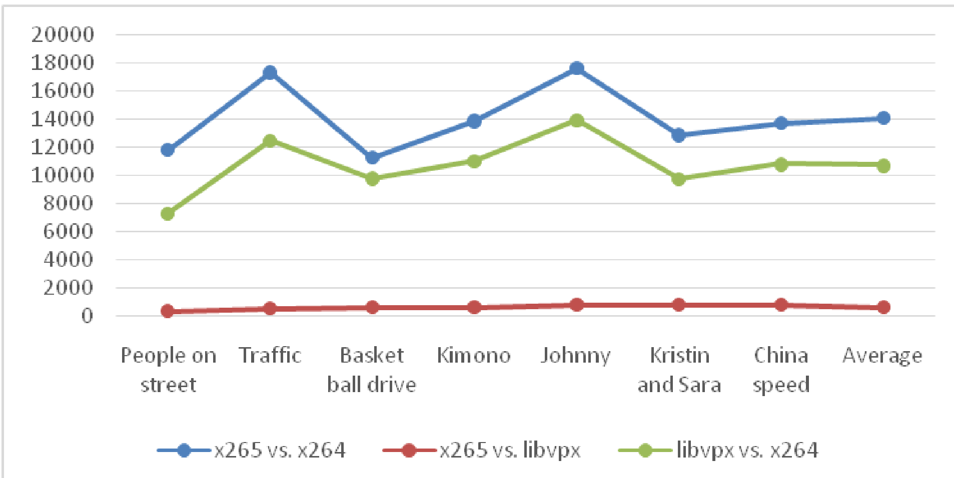

Figure 2. Time saving comparison between HEVC, H.264 and VP9

As shown in Figure 2. The x265 encoder consumed higher encoding time compared to x264 and libvpx. However, x264 required the lowest encoding time. The difference in encoding time between x264 and the other two codecs is higher than the difference between x265 and libvpx.

4.2 Subjective Experimental Results

The Bjontegaard of the Mean Opinion Score (MOS) results are shown in Table 7 and shown as BD-MOS. The BD-MOS can be calculated in the same way as BD-BR but MOS values instead of BR values.As it can be seen in the table, the variations between codecs are considered small. However, x265 outperformed other codecs performance and gained more user satisfaction than x264 and libvpx. The Libvpxencoder performed better than x264 and provided higher BD-MOS results. For the video sequences with high frame rates and resolutions (People on street and Traffic), the x265 provided almost similar improvement compared to the x264 and libvpx. However, the difference increased with lower frame rates and resolutions.

Table 7. Subjective comparison between HEVC, H.264 and VP9

	BD-MOS		
Sequence	**X265 vs. X264**	**X265 vs. Libvpx**	**Libvpx vs. x264**
People on street	52.3%	49.6%	18.4%
Traffic	72.6%	68.1%	26.8%
Basketball drive	71.7%	53.2%	23.6%
Kimono	64.1%	42.2%	17.2%
Johnny	58.5%	40.7%	15.7%
Kristin and Sara	50.4%	39.4%	12.3%
China speed	48.3%	36.6%	11.6%
Average	59.7%	47.1%	17.9%

As shown in Figure 3, the x265 provided the highest user satisfaction and outperformed x264 and libvpx. Furthermore, the codeclibvpx performed better than x264 and provided higher user satisfaction.

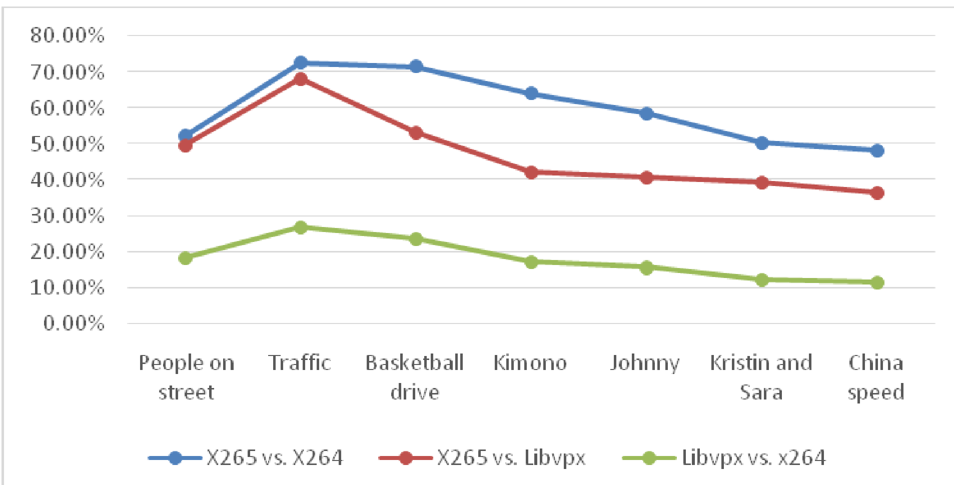

Figure 3. Subjective comparison between HEVC, H.264 and VP9

5. CONCLUSION

A performance evaluation of three state of the art video coding techniques was presented. The encoders H.265/MPEG-HEVC, VP9 and H.264/MPEG-AVC were evaluated using subjective and objective metrics. The experimental results showed that H.265/MPEG-HEVC outperformed VP9 and H.264/MPEG-AVC in the subjective tests. In addition, the VP9 outperformed H.264/MPEG-AVC in the subjective tests and provided better user satisfaction. However, H.265/MPEG-HEVC and VP9 required higher encoding time than H.264/MPEG-AVC.

REFERENCES

[1] T. Wiegand, G.J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC videocoding standard," Circuits and Systems for Video Technology, IEEE Transactions on , vol.13, no.7, pp.560-576, Jul. 2003.

[2] H.264/AVC Software Coordination, JM Reference Software, Online: http://iphome.hhi.de/suehring/tml

[3] Generic Coding of Moving Pictures and Associated Audio Information - Part 2: Video, 1994:ITU-T and ISO/IEC JTC 1.

[4] ITU-T, Recommendation H.265 (04/13), Series H: Audiovisual and Multimedia Systems, Infrastructure of audiovisual services – Coding of Moving Video, High Efficiency Video Coding, Online: http://www.itu.int/rec/T-REC-H.265-201304-I.

[5] J. 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)," Circuits and Systems for Video Technology, IEEE Transactions on , vol. 22, no.12, pp.1669-1684, Dec. 2012.

[6] M. Horowitz, F. Kossentini, N. Mahdi, S. Xu, H. Guermazi, H. Tmar, B. Li, G. J. Sullivan, J. Xu, "Informal subjective quality comparison of video compression performance of the HEVC and H.264/MPEG-4 AVC standards for low-delay applications," Proc. SPIE 8499, Applications of Digital Image Processing XXXV, 84990W, Oct. 15, 2012.

[7] WebM™: an open web media project, VP8 Encode Parameter Guide, 2-Pass Best Quality VBR Encoding, Online: http://www.webmproject.org/docs/encoder-parameters/#2-pass-best-qualityvbr-encoding

[8] J. Bankoski, P. Wilkins, X. Yaowu, "Technical overview of VP8, an open source video codec for the web," Multimedia and Expo (ICME), 2011 IEEE International Conference on, pp.1,6, 11-15 Jul. 2011.

[9] Chromium® open-source browser project, VP9 source code, Online:http://git.chromium.org/gitweb/?p=webm/libvpx.git;a=tree;f=vp9;hb=aaf61dfbcab414bfacc311501be17d191ff8506

[10] J. Bankoski, R. S. Bultje, A. Grange, Q. Gu, J. Han, J. Koleszar, D. Mukherjee, P. Wilkins, and Y. Xu, "Towards a next generation open-source video codec," Proc. SPIE 8666, Visual Information Processing and Communication IV, Feb. 21, 2013, pp. 1-13.

[11] HEVC Reference Software, Online: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware

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

[13] x264 free library/codec, 32-bit, 8-bit depth version r2334 (May 22, 2013), Online: http://www.divx-digest.com/software/x264.html

[14] x264 Settings, Online: http://mewiki.project357.com/wiki/X264_Settings

[15] G. Correa, P. Assuncao, L. Agostini, L.A. da Silva Cruz, "Performance and computational complexity assessment of High-Efficiency Video encoders," Circuits and Systems for Video Technology, IEEE Transactions on, vol.22, no.12, pp.1899-1909, Dec. 2012.

[16] Ronald Bultje, Google® Groups "WebM Discussion", Online: https://groups.google.com/a/webmproject.org/forum/?fromgroups#!topic/webmdiscuss/xopTll6KqGI

[17] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves", ITU-T Q.6/SG16 VCEG 13th Meeting, Document VCEG-M33, Austin, USA, Apr. 2001.

[18] International Telecommunication Union, (2002) "Methodology for the Subjective Assessment of the Quality of Television Pictures," Recommendation ITU-R BT. 500-11 ITU.

Author

Dr. Ahmad A. Mazharhas been a member of the College of Computing and Informaticsat Saudi Electronic University since 2015. He has more than ten years teaching experience. He received his Ph.D. in 2013 from De Montfort University, UK; his Master's degree in computer science from Al- Balqa' Applied University, Salt, Jordan; and his Bachelor's degree in computer science from Al-Zaytoonah University, Amman, Jordan.