To be published in Applications of Digital Image Processing XL, edited by Andrew G. Tescher, Proceedings of SPIE Vol. 10396 © 2017 SPIE

Performance Comparison of AV1, JEM, VP9, and HEVC Encoders

Dan Grois, Tung Nguyen, and Detlev Marpe
Video Coding & Analytics Department
Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute, Berlin, Germany
grois@ieee.org,{tung.nguyen,detlev.marpe}@hhi.fraunhofer.de

ABSTRACT

This work presents a performance evaluation of the current status of two distinct lines of development in future video coding technology: the so-called AV1 video codec of the industry-driven Alliance for Open Media (AOM) and the Joint Exploration Test Model (JEM), as developed and studied by the Joint Video Exploration Team (JVET) on Future Video Coding of ITU-T VCEG and ISO/IEC MPEG. As a reference, this study also includes reference encoders of the respective starting points of development, as given by the first encoder release of AV1/VP9 for the AOM-driven technology, and the HM reference encoder of the HEVC standard for the JVET activities. For a large variety of video sources ranging from UHD over HD to 360° content, the compression capability of the different video coding technology has been evaluated by using a Random Access setting along with the JVET common test conditions. As an outcome of this study, it was observed that the latest AV1 release achieved average bit-rate savings of ~17% relative to VP9 at the expense of a factor of ~117 in encoder run time. On the other hand, the latest JEM release provides an average bit-rate saving of ~30% relative to HM with a factor of ~10.5 in encoder run time. When directly comparing AV1 and JEM both for static quantization parameter settings, AV1 produces an average bit-rate overhead of more than 100% relative to JEM at the same objective reconstruction quality and, in addition, with a factor of ~2.7 in encoder run time. Even when operated in a two-pass rate-control mode, AV1 lags behind both the JEM and HM reference encoder with average bit-rate overheads of ~55% and ~9.5%, respectively, although the latter being configured along one-pass static quantization parameter settings.

Keywords: AV1, AOM, VP9, Future Video Coding, JEM, JVET, HEVC, H.265, HM, coding efficiency

1. INTRODUCTION

The amount of video traffic over the global communication network is drastically increasing every year. Currently, it is expected that Internet Protocol (IP) video traffic will be 82% of all consumer Internet traffic by 2021 [1]. One of the main reason for such a dramatic increase, is a high demand for high-resolution video content, particularly for the High Definition (HD) and Ultra-High Definition (UHD) content (the term UHD refers to both 3840x2160 (4K) or 7680x4320 (8K) resolutions in terms of luma samples).

In order to provide such high-quality content for a plurality of users, efficient video compression techniques should be employed. Currently, the High-Efficiency Video Coding (HEVC) standard, which is the latest standard developed by the Joint Collaborative Team on Video Coding (JCT-VC), is effectively deployed worldwide [2]. HEVC was established by both ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Pictures Expert Group (MPEG) in 2010, and its 1st version was approved by ITU-T as Recommendation H.265 [3] and by ISO/IEC as MPEG-H, Part 2 in 2013. Then, its 2nd version containing the Range Extensions (RExt) as well as the Scalable and Multiview Extensions (SHVC and MV-HEVC, respectively) was accomplished in 2014 [4]. Finally, the 3rd and 4th HEVC versions containing the 3D Video Coding Extensions (3D-HEVC) [5] and the Screen Content Coding Extension (HEVC-SCC) [6] were approved in 2015 and 2016, respectively.

However, more efficient video compression techniques are still desired, especially for streaming UHD video content, panorama video content from sport events, concerts, shows, etc., and 360° video content. Therefore, in order to fulfill this need, just two years after approval of the above-mentioned 1st HEVC version, the so-called Joint Video Exploration Team (JVET) on Future Video Coding of ITU-T VCEG and ISO/IEC MPEG was established with the aim to explore future video coding technologies beyond HEVC. Since then, the most promising explored future technologies have been integrated into the Joint Exploration Test Model (JEM) [7][8] [9]. The official standardization activities for the next video coding standard beyond HEVC are expected to be started in April 2018 after evaluation of the submissions to the Call for

Proposals (CfP) for future video coding technologies [10]. The next video coding standard is currently targeted to be finalized by the end of the year 2020.

In parallel with the above-mentioned standardization activities, a number of private companies have been developing their own video coding formats, such as, e.g., VP8 [11] or VP9 [12]. As is known, VP8 was developed by On2 Technologies, which was acquired by Google in 2009. The development of VP9, which was fully based on VP8, was accomplished in the middle of 2013. Performance comparisons of VP9 vs. H.265/MPEG-HEVC and H.264/MPEG-AVC codecs based on objective assessments were presented in detail in [13], [14], and [15], while a subjective assessment of VP9 vs. HEVC can be found in [16]. There were some additional activities related to the development of the so-called VP10 coding scheme [12] that was supposed to provide a significant gain over VP9; however, to the best of authors knowledge, the VP10 development has been suspended.

On the other hand, in 2015, the so-called Alliance for Open Media (AOM) was established [17][18], targeting to develop next-generation media formats with an especial emphasis on the development of a royalty-free video format. The first version of the AOM video format, so-called AV1, was released in April 2016, while the final version of AV1 has not yet been officially announced and will presumably not be available before mid of 2018. As declared by AOM, AV1 is targeting to provide a significant coding-efficiency gain over current state-of-the-art video codecs. However, presently there is a lot of confusion about the ability of AV1 to compete with other codecs, such as HEVC-based codecs, or even with JEM, which integrates future video coding tools beyond HEVC. In order to eliminate any such confusion, the authors of this paper provide detailed experimental results for the most recent available encoder realizations of the four video coding technologies AV1, VP9, HEVC, and JEM. While HM and JEM encoders are typically tested by using static quantization parameter (QP) configurations according to their common test conditions (CTCs) [9][20] as well as according to the CfP [10], the VP9 and AV1 encoders are evaluated by using two different approaches: 1) based on static QP configurations matching the above HM and JEM CTCs and their corresponding target bit-rates; and 2) utilizing their built-in multi-pass rate-control mechanisms. The detailed settings along with a discussion of the selected software implementations, the choice of coding parameters, and the corresponding evaluation setup are given in the course of this paper.

This paper is organized as follows. In Section 2, the selected representative encoders are introduced, while Section 3 contains a description of the test methodology and evaluation setup of the above-mentioned two different approaches: i.e., by using static QP configurations and by using multi-pass rate-control mechanisms for AV1 and VP9 coding schemes. Finally, the detailed experimental results are presented in Section 4, followed by the conclusions in Section 5.

2. SELECTED ENCODER IMPLEMENTATIONS

In this work, when selecting the representative encoders, it was kept in mind that the final release of AV1, if accomplished to be provided around mid-2018, will be targeting to compete not only with HEVC but also with future video coding technologies, such as presently explored by JEM [7]. Therefore, in this work, in addition to the HM encoder, the latest JEM encoder release is evaluated as a representative of the state-of-the-art in compression performance for a collection of future video coding tools under exploration. A brief overview of all selected representative encoders is presented below.

2.1 VP9 Encoder

To the best of authors knowledge, the first version of AV1 which was released by AOM in April 2016 [17], is fully based on VP9 (with all default tools enabled). This was also verified by the authors detailed analysis of the AV1 code repository [18]. Therefore, this version was selected by the authors for evaluating the VP9 coding scheme. As is known, the VP9 encoder [12] has a two-pass rate-control encoding option, which results in improved rate-distortion performance. This feature was enabled for VP9 as well as for its descendant AV1 in the multi-pass rate-control test case, as already mentioned above.

2.2 AV1 Encoder

For evaluating the AV1 encoder, the authors selected its latest release that was available in August 2017 [18]. It should be pointed out that this AV1 version is not the final one and further coding-efficiency improvements towards the completion of the official AV1 release are likely to be expected. It should also be noted that parallel processing in AV1 strictly relies on the usage of tiles, thereby implying a coding-efficiency penalty relative to the case where no tiles are used. As a consequence, multithreading in AV1 has been disabled in the experiments presented in Section 4 of this paper to not adversely affect its coding efficiency.

2.3 JVET JEM Encoder

The JEM encoder was selected for evaluating a collection of future video coding tools that have been implemented on top of HEVC [8]. Particularly, the authors used version JEM 6.0, which is one of the recent available versions. However, it should be pointed out that the collection of JEM coding tools as a whole is neither officially announced to serve as the starting point nor will it most likely be anywhere near the endpoint of the upcoming standardization activity towards the next-generation video coding standard. With this in mind, it is reasonable to expect that the final future video coding standard as a successor of HEVC will provide a compression capability that significantly exceeds that of JEM 6.0, in particular when compared for a given budget of encoder/decoder complexity.

2.4 H.265/MPEG-HEVC Reference Software Encoder

For H.265/MPEG-HEVC-based encoding, the HM reference software encoder [19] was selected as the main HEVC representative. It should be noted that the HM reference software is considered to be the most popular HEVC-based implementation for evaluating the coding-efficiency potential of all HEVC-conforming coding tools, even though there are numerous HEVC-compliant encoders that can provide better trade-offs between encoding speed and coding efficiency as, e.g., discussed and evaluated in [15]. In this paper, for conducting the corresponding HEVC performance evaluation, the authors selected HM version 16.15 [19] whose configuration can be easily adjusted to the corresponding JVET JEM CTC [9], particularly to match the corresponding GOP sizes and other encoding parameters.

3. TEST METHODOLOGY AND EVALUATION SETUP

The authors of this paper used very similar settings for all tested encoders, i.e., for AV1, VP9, JEM and HM reference software. While JEM and HM encoders were tested by using framewise fixed QP configurations in accordance with their corresponding CTC [9][20], the VP9 and AV1 coding schemes were tested by using two approaches: 1) by using framewise fixed QP configurations matching the above-mentioned HM and JEM CTCs; and 2) by using multi-pass rate-control mechanisms, as provided by the their corresponding encoders. The detailed test methodology and the evaluation setup are explained in detail below. Particularly, in Section 3.1, the VP9 and AV1 configurations are discussed, followed by the discussion of the JEM reference software configuration in Section 3.2. Further, in Section 3.3, the HM reference software configuration is presented. Finally, the employed Bjøntegaard-Delta Bit-Rate (BD-BR) metric is explained in Section 3.4.

3.1 VP9 and AV1 Configurations

In order to ensure a fair performance evaluation, the authors provided similar configuration settings for all tested encoders. Particularly, in Section 3.1.1 below, VP9 and AV1 encoders were set to corresponding fixed QP configurations with the rate control disabled similarly to other tested encoders. In addition, Section 3.1.2 provides the alternative configurations for VP9 and AV1 that allow a signal-adaptive choice of QPs within their built-in multi-pass rate control. However, it should be noted that the configuration settings in Section 3.1.2 are considered to be less fair with regard to HM and JEM encoders, since these encoders are one-pass encoders relying on framewise fixed/static QP settings.

3.1.1 Fixed QP Configurations for VP9 and AV1

The fixed QP configuration settings along with the specification of the downloaded software versions of VP9 and AV1 are presented in Table 1 below. The used configurations correspond to the latest recommended best-quality settings [12][18]. The QP range of VP9 and AV1 coding scheme was adjusted to fit the bit-rates produced by the JEM and HM reference software encoders. The parameters "min-gf-interval" and "max-gf-interval", which are related to 'golden' and 'alternate' reference frames (see Sect. 3.3), are configured equal to 16 since this choice corresponds to the temporal key-frame distance configuration of the HM/JEM encoders.

Table 1: Selected Fixed QP Configuration Settings for the VP9/AV1 Encoders.

	VP9 (1st AV1 version)	AV1
Git-Hash	b6724815f22876ca88f43b57dba09a555ef4e1b0 (Sep. 2016)	e784b3f2e6a40967c908ab585b12773172060020 (Aug. 2017)
Configuration	--best --tune=psnr --end-usage=q -b 10 --min-gf-interval=16 --max-gf-interval=16	

The reader is referred to [12] and [18] for obtaining more detailed information regarding the settings of Table 1. The intra-refresh interval, i.e., the periodic insertion of I-pictures at regular time intervals is controlled by the parameters "kf-min-dist" and "kf-max-dist" that are adjusted to be identical and sequence-specific matching the specification of the CTCs [9][20]. Also, it should be noted that the encoder configurations were tuned for best peak-signal-to-noise ration (PSNR) values. The quantization parameters are controlled by the parameters "min-q" and "max-q". By using different values for the two parameters, the encoders can independently choose a QP within the given range that may result in better compression efficiency. This is assumed to result in a similar effect to the hierarchical QP cascading that is employed for the HM/JEM encoders. The difference between the minimum and the maximum QP was set equal to eight, resulting in a similar range that is used for the hierarchical QP cascading in HM/JEM. Since the same HM/JEM QP values result in different bit-rates for VP9/AV1, different QP values were used to achieve a similar bit-rate range. These values were 27, 35, 46, and 55 for VP9/AV1 instead of 22, 27, 32, and 37 for HM/JEM.

3.1.2 Multi-Pass Rate-Control Configurations for VP9 and AV1

The multi-pass rate-control configuration settings for VP9/AV1 are presented in Table 2 below. The authors used the latest recommended best-quality settings [12][18].

Table 2: Selected Multi-Pass Rate-Control Configuration Settings for the VP9 and AV1 Encoders.

	VP9 (1st AV1 version)	AV1
Git-Hash	b6724815f22876ca88f43b57dba09a555ef4e1b0 (Sep. 2016)	e784b3f2e6a40967c908ab585b12773172060020 (Aug. 2017)
Configuration	--best --tune=psnr --end-usage=vbr --passes=2 -b 10 --min-gf-interval=16 --max-gf-interval=16 --min-q=0 --max-q=63	

The main difference relative to the fixed QP configuration is the "end-usage" parameter that is "vbr" instead of "q". Furthermore, the minimum quantization value is set to be the minimum allowed value, and similarly, the maximum quantization value is set to be the maximum allowed value. By such a way, the encoder obtains a maximum control for achieving the target bit-rate, which in turn allows the encoder a sequence-specific bit allocation (i.e., signal-adaptive QP setting) in such a way that the objective reconstruction quality is maximized.

3.2 JEM Reference Software Configuration

For the JEM reference software encoder, the Random Access (RA) configuration based on dyadic high-delay hierarchical B pictures was selected. The corresponding "encoder_randomaccess_jvet10.cfg" was used. The Group of Picture (GOP) size is set equal to 16 pictures, as given by the selected configuration. No coding tool related parameters have been changed. The choice of selecting the RA configuration was motivated by the fact that, firstly, this kind of the RA configuration is typically chosen for all broadcasting and streaming applications, and secondly, it provides in most cases better results in terms of coding efficiency compared to low delay configurations [14]. The intra-refresh interval was set to 32, 32, 48, and 64 pictures for video content with 24, 30, 50, and 60 fps, respectively. The hierarchical B pictures were encoded using a fixed QP increase of 1 (i.e., a quantization step size increase of around 12%) per frame between successive temporal levels. It is noted that the above test conditions were selected similarly to the test conditions presented in [9][10].

3.3 HM Reference Software Configuration

For the HM reference software encoder [19], similarly to the JVET JEM encoder, the RA configuration based on dyadic high-delay hierarchical B pictures was selected as well [20]. The corresponding "encoder_randomaccess_main10.cfg" was used with no further coding tool related parameters being changed. The Group of Picture (GOP) size is set equal to 16 pictures as described in the JVET JEM CTC [9], thereby employing one additional temporal layer relative to the GOP 8 structure. It should be noted that such an adjustment cannot easily be adapted to VP9 and AV1, since they do not include the concept of bi-predictive motion-compensation. A kind of alternative for bi-prediction is given in VP9/AV1 by the usage of so-called 'alternate' or 'golden' reference frames, which can be varied within a configuration by setting a minimum and maximum temporal distance. The intra-refresh intervals for HM were chosen to be exactly the same as for JEM, depending on the frame rate of the input video.

3.4 Bjøntegaard-Delta Bit-Rate Measurements

Similarly to the authors' previous work [13][14][15] the Bjøntegaard-Delta bit-rate (BD-BR) measurement method was used for the R-D performance assessment in order to calculate average bit-rate differences between R-D curves for the same distortion (e.g., for the same $PSNR_{YUV}$ values) [21][22]. It should be noted that negative BD-BR values indicate actual bit-rate savings.

The authors used R-D curves of the combined luma (Y) and chroma (U,V) components, while the combined $PSNR_{YUV}$ value were calculated as a weighted sum of the PSNR values per each picture of each individual component, i.e., of $PSNR_Y$, $PSNR_U$, and $PSNR_V$.

$$PSNR_{YUV} = (6 \cdot PSNR_Y + PSNR_U + PSNR_V) / 8 \quad (1)$$

As a result, using the combined $PSNR_{YUV}$ and bit-rate values as an input to the BD-BR measurement method enables to determine a single average difference in bit-rate that takes into account the reconstruction fidelity of both the luma and the two chroma components [21][22].

4. EXPERIMENTAL RESULTS

As explained in details in Section 3, the tests have been conducted in two different ways: 1) by using fixed quantization parameters (QP) for all tested codecs; and 2) by operating AV1 and VP9 in multi-pass rate-control mode. The detailed experimental results are presented in Section 4.1 and Section 4.2, respectively.

For obtaining all experimental results presented in this section, the corresponding test video sequences were selected according to both HEVC and JVET common test conditions (CTC) [9][20]. Additional sequences from the JVET JEM Call for Proposals (CfP) [10] were included. A special emphasis was put on video sequences having a 10-bit sample representation and UHD spatial resolution (particularly, the 4K resolution), as well as 360° video content. The corresponding test sequences are listed in Table 3 and Table 4.

Table 3: Test Sequences According to HEVC CTC [20].

	Class A	**Class B**	**Class C**	**Class D**	**Class E**
Resolutions	*2560×1600*	*1920×1080*	*832×480*	*416×240*	*1280×720*
Sequences	Traffic	Kimono	BasketballDrill	BasketballPass	FourPeople
	PeopleOnStreet	ParkScene	BQMall	BQSquare	Johnny
	Nebuta	Cactus	PartyScene	BlowingBubbles	KristenAndSara
	SteamLocomotive	BasketballDrive	RaceHorses	RaceHorses	

Regarding the additional sequences, Class A1 and A2 10-bit sequences are used together with additional newly introduced 10-bit video Classes A3 and B1, as listed in Table 4, while Class A3 contains 4K-resolution 10-bit sequences and Class B1 contains HD 10-bit sequences. It should be noted that classes having lower spatial resolutions were removed from the CfP [10].

Table 4: Additional Test Sequences, most of them are from the JVET CTC [9].

	Class A1	**Class A2**	**Class A3**	**Class B1**	**360°**
Resolutions	*3840x2160*	*3840x2160*	*3840x2160*	*1920×1080*	*4096×2048*
Sequences	Tango	CatRobot	BuildingHall	RitualDance	ChairliftRide
	Drums100	TrafficFlow	Crosswalk	Timelapse	Harbor
	CampfireParty	DaylightRoad	FoodMarket		KiteFlite
	ToddlerFountain	Rollercoaster	ParkRunning		SkateboardInLot
					Trolley

Moreover, due to the increased popularity, 360° video sequences (as listed in Table 4) were tested as well. The selected 360° sequences are taken also from the JEM CTC and are available in 8K resolution using the Equi-Rectangular Projection (ERP). According to the JVET CTC, a full evaluation of 360° sequences requires an extra software library, which provides software patches for HM and JEM, so that the 8K resolution sequences can be projected to a projection format of choice. When selecting ERP, the 8K input is simply downsampled to the 4K resolution prior to encoding. Then, the compression performance is evaluated by using PSNR at 4K resolution level. Other more meaningful metrics typically require additional input parameters, such as, e.g., a proper selection of the viewport. Such an analysis, however, would exceed the scope of this paper. Hence, the evaluation in this paper has been restricted to the traditional end-to-end PSNR measurement. It should also be noted that a similar library for evaluation of 360° sequences as specified by the JVET CTC is not publicly available for VP9/AV1.

4.1 Experimental Results for Fixed Quantization Parameters (QP)

Table 5 presents a summary of calculated BD-BR savings for JEM vs. VP9 as well as for JEM vs. AV1 per each Class A to 360°, where negative BD-BR values indicate bit-rate savings (in contrast to positive values, which indicate the required overhead in bit-rate to achieve the same $PSNR_{YUV}$ values). It should be noted that first version of AV1, which was released by AOM in April 2016 [17] and which is denoted in this work as VP9, does not support 10-bit encoding. Therefore, the encoding of Class A1, A2, A3 and B1 sequences in this work cannot be evaluated by using the VP9 coding scheme and the two 10-bit sequences "Nebuta" and "SteamLocomotive" of Class A were left out.

Table 5: Summary of Results of JEM Encoder vs. VP9 and AV1 Encoders for the Fixed QP Configuration.

Sequence class	***JEM* relative to**		*Missing results for AV1*
	VP9	***AV1***	
A	-58.0%	-47.7%	Nebuta
A1		-38.6%	
A2		-56.8%	
A3		-53.9%	ParkRunning
B	-62.9%	-52.4%	
B1		-51.2%	
C	-60.4%	-51.6%	
D	-59.6%	-51.0%	
E	-69.4%	-60.5%	
360°	-61.6%	-51.6%	ChairliftRide
Averaged BD-BR	**-62.0%**	**-51.4%**	
Factor of *JEM* Encoder Run Time	**36.99x**	**0.37x**	

As seen from Table 5, the average BD-BR savings of the JEM reference encoder relative to VP9 and AV1 encoders are 62.0%, and 51.4%, respectively. There are no outliers in terms of compression efficiency performance, i.e., the BD-BR values across different classes are fairly close to each other. In addition, the increase in encoding computational complexity (in terms of encoder run time) of the tested AV1 encoder compared to the JEM 6.0 encoder is a factor of ~2.7.

Further, Table 6 below presents a summary of the BD-BR savings of HM over both VP9 and AV1 encoders. As seen from Table 6, the average BD-BR savings of the HM reference encoder relative to VP9 and AV1 codecs are 46.6%, and 30.6%, respectively.

Table 6: Summary of Results of HM Encoder vs. VP9 and AV1 Encoders for the Fixed QP Configuration.

Sequence class	***HM* relative to**		*Missing results for AV1*
	VP9	***AV1***	
A	-39.9%	-24.3%	Nebuta
A1		-12.4%	
A2		-30.9%	
A3		-31.4%	ParkRunning
B	-47.8%	-33.4%	
B1		-31.8%	
C	-44.9%	-33.2%	
D	-44.5%	-33.0%	
E	-54.7%	-42.5%	
360°	-46.4%	-34.8%	ChairliftRide
Averaged BD-BR	**-46.6%**	**-30.6%**	
Factor of *HM* Encoder Run Time	**3.44x**	**0.04x**	

The increase in the encoder run time for HM relative to the VP9 encoder is about a factor of 3.4; on the other hand, the fully-fledged HM reference software implementation encodes the above tabulated classes of video sequences about 25 times faster on average than the tested AV1 encoder implementation.

Table 7 provides a summary of the BD-BR performance evaluation (per each class) of the JEM vs. HM reference encoders. As seen from Table 7, the JEM reference software provides bit-rate savings of 29.8% compared to HM, while its encoding complexity in terms of encoder run time is about 11 times higher than that of HM.

Table 7: Summary of Results of JEM vs. HM Encoders and of AV1 vs. VP9 Encoders for the Fixed QP Configuration.

Sequence class	***JEM* relative to *HM***	***AV1* relative to *VP9***	***Missing results for AV1***
A	-28.3%	-15.4%	
A1	-31.2%		
A2	-38.0%		
A3	-31.1%		
B	-28.8%	-19.5%	
B1	-28.2%		
C	-28.1%	-15.8%	
D	-27.4%	-15.7%	
E	-31.5%	-18.2%	
360°	-26.0%	-16.8%	ChairliftRide
Averaged BD-BR	**-29.8%**	**-17.1%**	
Encoder Run Time Factor	**10.51x**	**117.55x**	

Furthermore, Table 7 provides a summary of the performance evaluation of the AV1 vs. VP9 encoders. AV1 outperforms VP9 by 17.1% in bit-rate savings. However, the AV1 encoder spends much more time to encode the classes of 8-bit video sequences as presented in Table 7; particularly, the factor in encoder run time of AV1 relative to VP9 is about 117. This means that the quite substantial average bit-rate savings of 17.1% for AV1 relative to VP9 requires a huge amount of increase in computational complexity by a factor of more than 100 in encoder run time.

Table 8: Summarized BD-BR Experimental Results of All Tested Encoders for Fixed QP Configuration[1].

test candidate \ anchor	**AV1**	**JEM**	**VP9**	**HM**
AV1		111.8%	-17.1%	47.7%
JEM	-51.4%		-62.0%	-29.8%
VP9	21.0%	173.7%		92.5%
HM	-30.6%	43.4%	-46.6%	

To summarize the BD-BR performance of all tested encoders, Table 8 provides an overview of the outcome of the conducted experiments. As shown in Table 8, AV1 produces a bit-rate overhead of 47.7% compared to the HEVC-based encoder. In addition, in order to achieve the same average $PSNR_{YUV}$ values of HM, when using the VP9 encoder, a BD-BR overhead of 92.5% is required. When compared with JEM, AV1 has a significant bit-rate overhead of 111.8%, while the AV1 encoding is 2.7 times slower than the JEM encoding. In addition, the HM reference encoder has a coding gain of 30.6% over AV1, while the HM encoding is 25 times faster than that of AV1. On the other hand, VP9 encoding is 3.4

[1] Due to the fact that the fitting of R-D curves slightly differs when fitting the R-D curve of one encoder to another and vice versa, the product (100 + b1)(100 + b2) for each pair (b1, b2) of corresponding BD-BR values (e.g., JEM vs. AV1 and AV1 vs. JEM) is only approximately equal to 10.000

times faster than that of HM, whereas the bit-rate savings of HM vs. VP9 are 46.6%. When comparing AV1 and VP9, the bit-rate overhead of VP9 over the tested AV1 encoder is 21.0%, while the increase in encoder run time for AV1 amounts to a factor of 117 compared to that of VP9. The bit-rate savings of JEM relative to HM are 29.8%, while HM encoding is about 11 times faster than that of JEM.

4.2 Experimental Results for Multi-Pass Rate-Control

Table 9 presents a summary of calculated BD-BR savings for JEM vs. VP9 as well as for JEM vs. AV1 for the case where the AV1 and VP9 encoders were evaluated with varying QPs within the multi-pass (i.e., 2-pass) rate-control scheme. It should be noted that for JEM (and HM) the same encoding results as in Section 4.1 are used. As already noted in Section 4.1, since VP9 does not support 10-bit encoding, the encoding of Class A1, A2, A3 and B1 sequences have been omitted for VP9.

Table 9: Summary of Results of JEM vs. VP9 and AV1 Encoders for the Rate-Control Configuration.

Sequence class	***JEM* relative to**		***Missing results for AV1***
	VP9	***AV1***	
A	-47.1%	-31.2%	Nebuta
A1		-31.9%	ToddlerFountain
A2		-39.8%	CatRobot
A3		-41.6%	FoodMarket, ParkRunning
B	-48.7%	-37.8%	
B1		-25.1%	
C	-42.6%	-29.0%	
D	-43.6%	-31.1%	
E	-54.1%	-39.7%	
360°	-48.8%	-38.5%	
Averaged BD-BR	**-47.3%**	**-34.8%**	
Factor of *JEM* Encoder Run Time	**16.64x**	**0.52x**	

As seen from Table 9, in case when VP9 and AV1 encoders are tested under the multi-pass rate-control settings, the average BD-BR savings of the JEM encoder are 47.3%, and 34.8%, respectively. In addition, compared to AV1, the JEM bit-rate savings are achieved with a decrease of a factor of ~1.9 in encoder run time.

Further, Table 10 below presents a summary of the BD-BR savings of HM over both VP9 and AV1 encoders under the multi-pass rate-control settings. As seen from Table 10, the average BD-BR savings of the HM reference encoder relative to VP9 and AV1 encoder are 27.0%, and 7.8%, respectively. Note that the average AV1 encoder run time is a factor of 20 times larger than that of HM, which in turn is a factor of 1.55 larger than the average VP9 encoder run time.

Table 10: Summary of Results of HM vs. VP9 and AV1 Encoders for the Rate-Control Configuration.

Sequence class	***HM* relative to**		***Missing results for AV1***
	VP9	***AV1***	
A	-25.3%	-2.1%	Nebuta
A1		4.6%	ToddlerFoundtain
A2		-7.6%	CatRobot
A3		-13.4%	FoodMarket, ParkRunning
B	-29.0%	-12.8%	
B1		2.7%	
C	-20.6%	-2.7%	
D	-22.8%	-6.6%	
E	-33.2%	-14.1%	
360°	-30.5%	-17.2%	
Averaged BD-BR	**-27.0%**	**-7.8%**	
Factor of *HM* Encoder Run Time	**1.55x**	**0.05x**	

Table 11 provides a summary of the BD-BR performance evaluation of the AV1 and VP9 coding schemes for the rate-control configurations. AV1 outperforms VP9 by 20.0% in terms of bit-rate savings, while it takes the AV1 encoder more than 35 times longer than the VP9 encoder to encode all 8-bit sequences, as summarized in Table 11.

Table 11: Summary of Results of AV1 vs. VP9 Encoders for the Rate-Control Configuration (8-bit sequences).

Sequence class	***AV1* relative to *VP9***
A	-18.4%
B	-29.0%
C	-18.0%
D	-16.5%
E	-20.1%
360°	-16.0%
Averaged BD-BR	**-20.0%**
Encoder Run Time Factor	**35.76x**

Table 12: Summarized BD-BR Experimental Results of All Tested Encoders for the Rate-Control Configuration.

		anchor			
		AV1	**JEM**	**VP9**	**HM**
test candidate	**AV1**		55.0%	-20.0%	9.5%
	JEM	-34.8%		-47.3%	
	VP9	28.5%	92.2%		37.9%
	HM	-7.8%		-27.0%	

To summarize the BD-BR performance of all tested encoders, Table 12 provides an overview of the outcome of the conducted experiments. As shown in Table 12, AV1 still produces a bit-rate overhead of 55.0% and 9.5% compared to JEM and HM encoders, respectively, despite the fact that both JEM and HM are tested only under the framewise fixed QP regime. Note that due to the same reason, the corresponding fields of JEM vs. HM and vice-versa in Table 12 are grayed out. Table 12 also shows that the HM reference encoder still provides coding gains of 7.8% and 27.0% over AV1 and VP9, respectively, even though AV1 and VP9 are operated along the multi-pass rate-control configuration settings.

The gap between the results when using fixed QP and rate control for both VP9 and AV1 is significant, as a comparison of the BD-BR values between Table 8 and Table 12 reveals. As an indication of the underlying behavior, Figure 1 and Figure 2 show plots of time series for three different coded versions of the sequence “BQTerrace.” Figure 1 shows the luma (Y) PSNR for each reconstructed frame, while Figure 2 shows the corresponding bits per frame in megabits (Mbit). The plot abbreviated as “HM” depicts the HM bitstream, the plot “AV1-RC” depicts the AV1 bitstream generated by using rate control, and the plot “AV1-QP” depicts the AV1 bitstream generated by using a fixed QP setting. Note that all three bitstreams correspond to the same target bit rate. For the HM bitstream (as shown by the black curve), a repeating pattern consisting of quite regular peaks in both PSNR and bits per frame can be observed. This pattern follows the employed hierarchical QP cascading structure of the HM encoder along the temporal (frame number) axis with the highest peaks generated by the key frames.

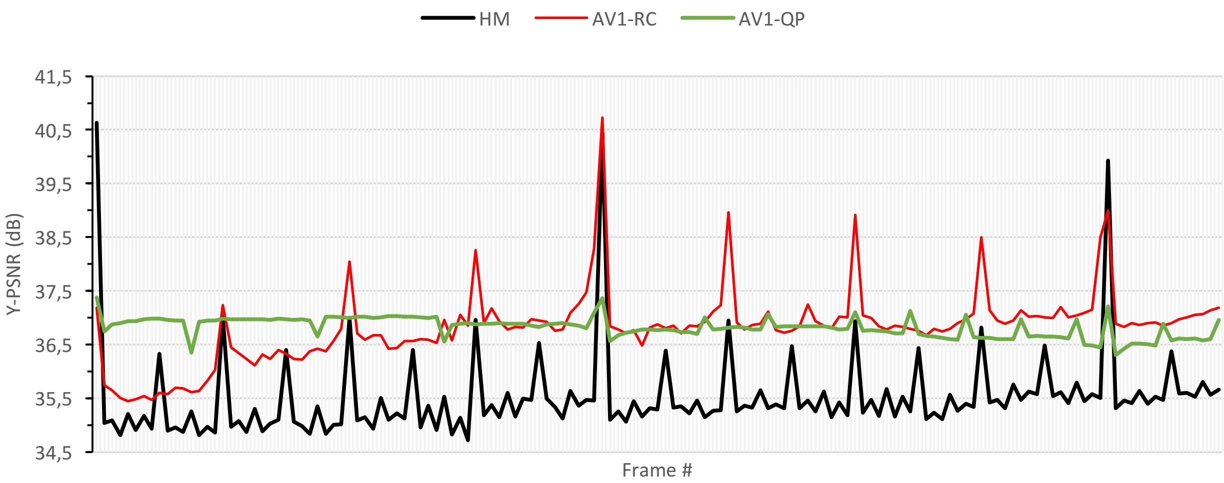

Figure 1: Y-PSNR (in dB) for the different reconstructions of HM using fixed QP ("HM"), AV1 using rate control ("AV1-RC"), and AV1 using fixed QP ("AV1-QP") plotted against the frame index for the "BQTerrace" video sequence.

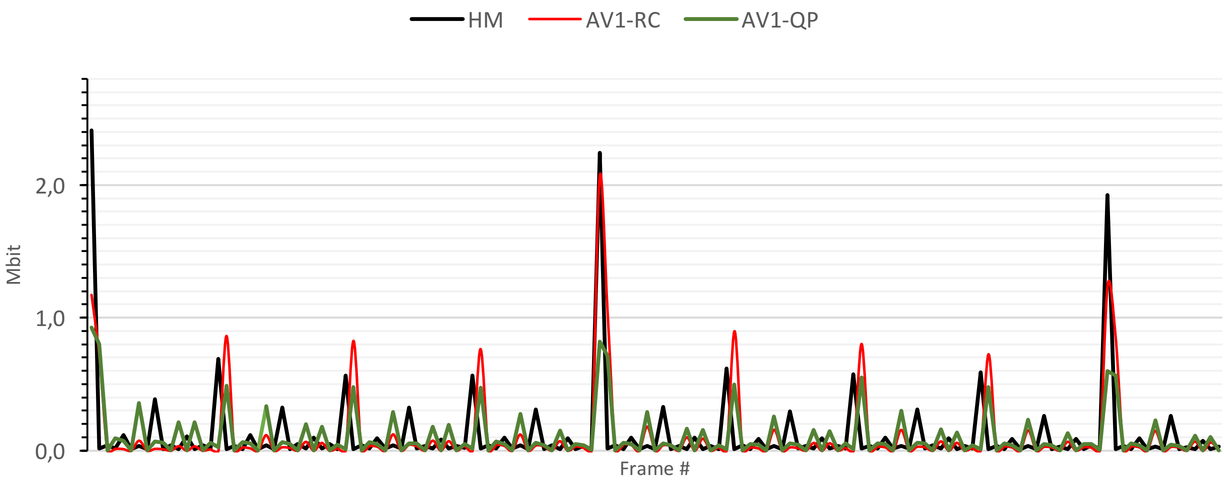

Figure 2: Bits per frame (in Mbit) for the different bitstreams belonging to the reconstructions, as shown in Figure 1: HM bitstream using fixed QP ("HM"), AV1 bitstream using rate control ("AV1-RC"), and AV1 bitstream using fixed QP ("AV1-QP").

In contrast to the plots of the HM bitstream, the plots of the AV1 bitstream for fixed QP (green curve) show significantly less variance along the temporal (frame #) axis, both in luma PSNR (Figure 1) and bits per frame (Figure 2). This behavior indicates that the AV1 encoder does not properly exploit the freedom in selecting the optimal QP for each frame within the given bounds of the parameters "min-q" and "max-q", as discussed in Section 3.1.1. However, when operated in rate-control (RC) mode, as depicted by the red curves, the variance of the output of the AV1 encoder both in terms of allocated bits and Y-PSNR per frame is considerably increased. With this behavior, the AV1-RC encoder seems to compensate for the lack of hierarchical B pictures and the corresponding quantization strategy of the HM encoder, and thus resulting in an overall improved rate-distortion performance. Note that, even though it seems to be counterintuitive at first sight, the AV1 encoder produces skipped frames represented with a (nearly) vanishing number of bits, as can be seen from the dips in the green and red curves in Figure 2. These skipped frames result from the use of non-displayed alternate reference frames, which are packed together with ordinary neighboring motion-compensated frames by means of so-called superframes that, on the other hand, can be identified as the peaks in between two key frame peaks of the green and red curves in Figure 2.

5. CONCLUSION

An extensive performance comparison of AV1, VP9, JEM, and HEVC encoders was presented and discussed in detail, while the evaluation was done with both fixed QP and multi-pass rate-control configurations. According to the experimental results, the coding efficiency of AV1 and VP9 was shown to be significantly inferior to both HM and JEM encoders. Particularly, for the fixed QP configuration, the AV1 coding scheme produced an average bit-rate overhead of 47.7% and 111.8%, at the same objective quality, relative to HM and JEM encoders, respectively. In addition, the JEM encoding was found to be almost three times faster than the AV1 encoding. Further, the HM reference encoder provided average bit-rate savings of 30.6% relative to AV1 with an average encoding speed that was 25 times faster than that of AV1. In comparison to VP9, HM encoding was found to be 3.4 times slower but more compression efficient with a significant bit-rate saving of 46.6%. In turn, the bit-rate overhead for VP9 over the recently available AV1 coding scheme was on average 21.0%, while the AV1 encoding time was measured to be about 114 times the VP9 encoding time. On the other hand, for the multi-pass rate-control configuration, AV1 and VP9 still showed some significant bit-rate overheads of 55.0% and 92.2%, respectively, compared to JEM, in spite of the fact that JEM was tested using the one-pass fixed QP settings. Finally, when comparing AV1 and VP9 with the same multi-pass rate-control enabled to the HM reference encoder (that similarly to JEM was tested using the one-pass fixed QP settings), the measurements indicated bit-rate overheads of 9.5% and 37.9%, respectively.

REFERENCES

[1] Cisco Visual Networking Index: Forecast and Methodology, 2016–2021", <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html>, Cisco Systems Inc., Jun. 2017.
[2] G. Sullivan, "Deployment status of the HEVC standard", document JCTVC-AB0020 of JCT-VC, Torino, IT, July 2017.
[3] 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.
[4] ITU-T, Recommendation H.265 (10/14), Series H: Audiovisual and Multimedia Systems, Infrastructure of audiovisual services – Coding of Moving Video, High Efficiency Video Coding.
[5] ITU-T, Recommendation H.265 (04/15), Series H: Audiovisual and Multimedia Systems, Infrastructure of audiovisual services – Coding of Moving Video, High Efficiency Video Coding, <http://www.itu.int/rec/T-REC-H.265-201504-I>.
[6] ITU-T, Recommendation H.265 (12/16), Series H: Audiovisual and Multimedia Systems, Infrastructure of audiovisual services – Coding of Moving Video, High Efficiency Video Coding.
[7] JVET JEM software <https://jvet.hhi.fraunhofer.de/>.
[8] J. Chen, E. Alshina, G. Sullivan, J.-R. Ohm, "Algorithm Description of Joint Exploration Test Model 6 (JEM 6)," document JVET-F1001, Hobart, AU, April 2016.
[9] K. Suehring, and X. Li, "JVET common test conditions and software reference configurations," document JVET-B1010 of JVET, San Diego, USA, Feb. 2016.
[10] A. Segall, V. Baroncini, J. Boyce, J. Chen, and T. Suzuki, "Joint Call for Proposals on Video Compression with Capability beyond HEVC," document JVET-H1002 of JVET, Macao, CN, Oct. 2017.
[11] WebM: an open web media project, VP8 Encode Parameter Guide <http://www.webmproject.org/docs/encoder-parameters>.
[12] Chromium open-source browser project, VP9 source code, <http://git.chromium.org/gitweb/?p=webm/libvpx.git;a=tree;f=vp9;hb=aaf61dfbcab414bfacc3171501be17d191ff8506>.
[13] D. Grois, D. Marpe, A. Mulayoff, B. Itzhaky, and O. Hadar, "Performance comparison of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC encoders," Picture Coding Symposium (PCS), 2013, pp.394-397, 8-11 Dec. 2013.
[14] D. Grois, D. Marpe, T. Nguyen, and O. Hadar, "Comparative Assessment of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC Encoders for Low-Delay Video Applications", Proc. SPIE Vol. 9217, Sept. 2014.
[15] D. Grois, T. Nguyen, and D. Marpe, "Coding Efficiency Comparison of AV1/VP9, H.265/MPEG-HEVC, and H.264/MPEG-AVC Encoders," Picture Coding Symposium (PCS), Dec. 2016.

[16] M. Řeřábek, and T. Ebrahimi, "Comparison of compression efficiency between HEVC/H.265 and VP9 based on subjective assessments", Proc. SPIE Vol. 9217, Sept. 2014.
[17] Alliance for Open Media, Press Release <http://aomedia.org/press-release/>.
[18] Git repositories on aomedia <https://aomedia.googlesource.com/>.
[19] JCT-VC HM software <https://hevc.hhi.fraunhofer.de/>.
[20] F. Bossen, "Common HM test conditions and software reference configurations," document JCTVC-L1100 of JCT-VC, Geneva, CH, Jan. 2013.
[21] 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)," IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, no.12, pp.1669-1684, Dec. 2012.
[22] 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.