SERBIAN JOURNAL OF ELECTRICAL ENGINEERING
Vol. 18, No. 2, June 2021, 139-154

UDC: 621.397:004.9 **DOI: https://doi.org/10.2298/SJEE2102139M**

Video Coding and Constant Quality Evaluation Using 4k aomenc-AV1 and rav1e-AV1 Formats

Milan Milivojević[1], Dragi Dujković[1], Ana Gavrovska[1]

Abstract: It is much expected from the relatively novel, open, royalty-free AV1 (Alliance for Open Media (AOMedia) Video 1) standard. At this moment, there are many new variants of AV1 format. It is designed for efficient video internet delivery and high-quality video transmission. AV1 is recognized as Google's VP9 format successor. One of the reference tools used so far for testing AV1 is libaom-AV1. Nevertheless, due to its time-consuming performance, there are now different available standalone solutions for experimental analysis. Here, one such solution AOMedia's standalone aomenc (aomenc-AV1) is tested in order to analyze quality assessment based on constant quality constraint factor. Three different metrics are calculated for various 4k video content of the same frame rate. Moreover, rav1e implementation was tested for the same visual data, where rav1e-AV1 also represents an AV1 video encoder, which is considered reliable and suitable in most cases, where libaom is not applicable. In this paper, the comparison results between aomenc-AV1 and rav1e-AV1 are shown.

Keywords: Video codecs, 4k/UHD, Constant Quality, aomenc-AV1, libaom-AV1, rav1e-AV1, Bitstream.

1 Introduction

It is well known that by 2022, more than 80% of global internet will be dedicated to video [1 – 2]. There are many on-going projects related to video delivery over internet [3]. Some of them are oriented towards high-resolution data streaming. On the other hand, the high-resolution video usually means premium content for content providers. The general focus is to develop new technologies, and new coding and compression standards, which will enable efficient video internet delivery.

The term "4k" refers to horizontal resolutions of around 4000 pixels. Typical example where 4k or UHD (Ultra-High Definition) content and efficient internet delivery is of importance is OTT (Over-the-Top) streaming. OTT providers like Netflix, Hulu, and many others may benefit from these new solutions. Coding

[1]School of Electrical Engineering, University of Belgrade, Bulevar kralja Aleksandra 73, 11120 Belgrade, Serbia; E-mails: msmilance@etf.bg.ac.rs; dragi@etf.bg.ac.rs; anaga777@etf.bg.ac.rs (anaga777@gmail.com)

should be available for different smart and consumer electronics devices, as well. Standard Television broadcasting can be improved.

There are several main focuses in video technology nowadays [4, 5]. One of them is to make a significant increase of viewers' expectations. In other words, besides high Quality of Service (QoS), Quality of Experience (QoE) has become extremely relevant for future development of ICT (Information and Communication Technologies). Innovative solutions in coding are needed for fast delivery of higher quality (UHD quality) immersive content to the viewers. So, the focus is to make innovations directly available for the market, providing fast improvements from both production and services standpoints. Quality of Business and returns of investments are also important, so management of possible risks and costs should be improved.

One of the most common standards used for video encoding and compression tasks is H.264/AVC (Advanced Video Coding). Even though, it is a relatively older solution (from about 2004) it is still one of the most popular. In 2013, another standard giving improved results has been proposed. It is H.265/HEVC (High Efficiency Video Coding) [6]. Besides HEVC, another standard called VP9 has become popular. It was developed by Google and is widely used on video platforms as open, royalty-free solution. Due to the need of higher efficiency, the video content providers joined their forces by forming AOMedia (Alliance for Open Media) [3, 7-9]. Their goal has been to develop a new open, royalty-free format like VP9 but with more success in coding and compression tasks. In 2018 the first release of AV1 (AOMedia Video 1 codec) format has been developed for the purpose of video delivery over internet.

Nowadays, there are many different AV1 formats [3]. The reference for experiments with the new format is libaom-AV1 [10, 11]. This reference can be used for analysis for comparison with AVC, HEVC or VP9. One of the main disadvantages is time needed for (re-)encoding, as tested in [12]. Due to this, new implementations that can be applied for AV1 general usage are tested here. In this paper, implementations, such as aomenc-AV1 and rav1e-AV1 are considered [13, 14].

The paper is organized as follows. The introduction is followed by Section 2, where AV1 (AOMedia Video 1) format is briefly explained. Main characteristics are presented, and several AV1 approaches are mentioned. Two alternatives to reference libaom software tool are selected, such as aomenc-AV1 and rav1e-AV1. The 4k video dataset and the simulation details are given in Section 3. Here, the focus is on aomenc-AV1 and rav1e-AV1 software comparison. In Section 4 the obtained experimental results for different quality factors are presented. The conclusions are given in Section 5.

2 AV1 format, libaom-AV1/aomenc-AV1 and rav1e-AV1 coding

AV1 (AOMedia Video 1) has been launched by AOMedia (Alliance for Open Media), which members (Amazon, Google, Cisco, Intel, Microsoft, Netlix and many others) united in developing the new format [3]. The AV1 format is expected to give not just satisfying results in comparison to VP9 or HEVC (High Efficiency Video Coding), and efficiently work with UHD (Ultra-High Definition) data, but to be useful in many implementation which include video delivery over internet [6, 7].

As well known Google's VP9, AV1 is considered as open, royalty-free standard. AOMedia has to deal with both software and hardware coding issues. In the initial releases, like libaom-AV1 [11], AV1 has only been considered from the software standpoint. Libaom is considered the reference software implementation, including encoder and decoder. Libaom has been a relatively time-consuming solution, but encoding and decoding solution has been improved in the meantime. Aomenc has been the encoder, but further work and optimizations made new tools which has been enhanced the aomenc-AV1 implementations. Thus, aomenc-AV1 can also be understood as a kind of a reference approach.

Already has been announced that the one AV1 solution is not enough since there is a need for variety of different implementations, where some of them are propriety solutions like Cisco AV1 for Webex teleconference implementation, where minimization for latency in the case of teleconferencing has been the primary goal. There are different propriety encoders like: Eve-AV1 from Two Oriols, Aurora (wzAV1) from Visioular, and many others, but there are also the non-propriety ones.

AVIF (AV1 Image File Format) is a new image format, which can be derived from the keyframes of AV1 format, where Chrome has already supported this format.

Netflix and Intel has started their AV1 project for developing SVT-AV1 (SVT - Scalable Video Technology) method, primarily focusing on the encoding task. Specific aim has been set in this case to develop open-source codec (encoder and decoder) that can be useful for OTT services and VoD (Video on Demand). Particularly of interest are Intel Xeon processors and optimization of the solution on such hardware for the applications on data center servers. It can be seen that video platforms, such as Youtube, has already implemented AV1 format.

Besides specific implementations, rav1e-AV1 is encoder considered for general usage. It is similar to aomenc-AV1, but the goal in this case has been to start from the most simple and fast solution and to increase the efficiency compared to the reference approach. It is called the fastest and the safest AV1 version designed to cover all common cases, and to be suitable when the libaom is not adequate for its time-consuming approach.

Dav1d is an open-source AV1 decoder developed by VideoLAN and FFmpeg community. It is sponsored by the Alliance for Open media, and represents cross-platform decoder with a focus on speed and correctness. This is a high-speed decoder for most platforms, in order to deal with hardware issues, and for low CPU (Central Processing Unit) processing. Google has been developing libgav1 solution, as their own AV1 decoder.

In Fig. 1 the general architecture for AV1 is illustrated. From the architecture standpoint it resembles HEVC approach. Google's VP9 standard is considered as predecessor, which can be seen through tracing experiments [12]. Future implementations should balance software and hardware possibilities. Open benchmarking is possible, mostly for 1080p (p – progressive scan) and 4k data [15].

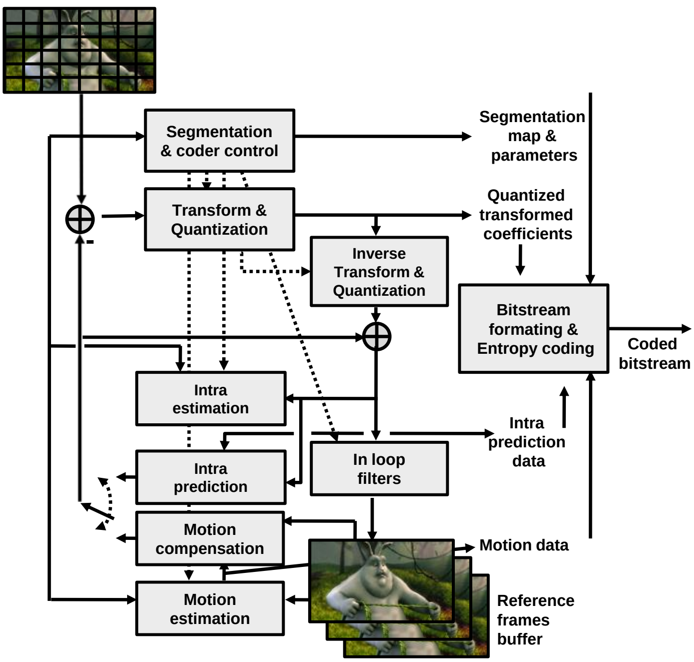

Fig. 1 – *General AV1 (Alliance for Open Media Video 1) architecture* [2].

Improvements in general AV1 format are numerous. The goal is the development of video coding format for optimized high-performance tasks [7, 12]. One should have in mind that a software improvement in coding architectures often means high computational cost. That is confirmed by the reference software libaom. Libaom is reference AV1 codec (coder-encoder), and it has enabled the initial insight into the AV1 advatages.

Gains are obtained by different approaches which are all part of one encoder or decoder solution [2]. One such approach that enables higher performance is using superblocks. The superblocks are of 128×128 pixels. They are introduced for the first time in AV1 format. Also, recursive partitioning is applied, as well as several partition approaches, as illustrated in Fig. 2. Hierarchy has been a part of AV1 approach, as well as recursive techniques.

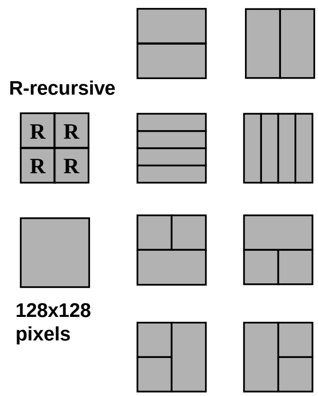

Fig. 2 – *Superblock partitioning introduced in AV1.*

In both space and time motion prediction has been improved with AV1. Motion vectors uses superblocks of 128×128 pixels, and smaller ones like 64×64 pixels available in VP9. Eight intra-prediction directional modes are applied.

Code words of ten or twelve bits are expected in comparison to standard eight bit depth. Rectangular DCT (Discrete Cosine Transform) and asymmetric DST (Discrete Sine Transform) are used in AV1 coding format. New quantization parameters and filtering techniques are adopted [2, 3, 7]. Delivering high-performance coding solution implies different algorithms, making rapid progress, and fast delivering of the techniques to the public.

There are two main approaches in experimental analysis of different coding solutions. One is to define constant video output quality, and to perform coding according to constant quality factor. Another approach is to make different constraints for bitrate. In this paper, the first approach for coding experimental analysis has been applied.

In [7] it is showed that libaom performs better than in the case of HEVC. About 43% improvement is reported for PSNR (Peak Signal-to-Noise ratio). In [16] 4k video traffic has been analyzed using prediction models, where the sequences were encoded using HEVC. In [17] the traffic variability has been analyzed. In our previous work [12] the reference libaom was applied for typical test 4k video sequences. The results in [12] are compared to traces obtained using VP9 and HEVC codecs implemented using ffmpeg. Constant quality settings are used for two 4k samples and four factors (20, 24, 30, 34). Two-norm evaluation is performed on trace sequences. The relative estimation showed that in the case of libaom application, smaller difference is obtained for lower quality for values 30, 34, compared to VP9. From the standpoint of frames, libaom can be differentiated from the size standpoint. It is noticed that there is a possibility to distinguish two groups, relatively small and relatively high sample values of traces. This shows higher control of the AV1 coding. In comparison to this, VP9 trace sequences showed more natural behaviour. Two-norm has been only comparable between AV1 and VP9, since HEVC has also B (bidirectional predicted) frames besides I (intra-coded) and P (predicted) type frames. Also, time needed for coding is measured confirming the libaom as a time-consuming approach. It is showed that about fifty times more minutes was needed for coding using libaom in some cases, compared to VP9 or HEVC. At the moment of writing this paper, new optimizations have been available, and aomenc-AV1 implementation. Experimenal analysis using recent SVT-AV1 implementation and 4k has been analyzed in our recent work [18].

In this paper rav1e-AV1 and the reference aomenc-AV1 tools are analyzed [13, 14].

3 Experimental Analysis

Experimental analysis is performed in the context of constant quality factor, named CQ (Constant Quality) or just Q (Quality). In the case of constant quality experiments, the goal is to achieve certain visual quality level without specific bitrate settings. Factor Q is selected in the experiment from a set of values (5, 10, 20, 30, 40, 50, 60), where the highest value means the lowest quality, and the lowest one corresponds to the highest video output.

Six mp4 video test inputs are used for experimental analysis, where mp4 represents the well known multimedia container MPEG-4 Part 14 (MPEG - Moving Picture Experts Group) [19]. Each of the six video sequences can be considered 4k (mostly 4k UHD (Ultra-High Definition) or UHD-1) and of low frame rate (23.98 frames per second (fps)) Details regarding the test video inputs are given in Table 1. The configuration of the processor used for tests here is: AMD Ryzen 5 4600H 4 GHz, with Graphics adapter NVIDIA GeForce GTX 1650 Ti Mobile – 4096 MB.

The initial tests presented in [12], showed reference tool libaom-AV1 as extremely time-consuming for tests. This is the reason why aomenc-AV1 was selected instead of libaom-AV1 in this paper.

Table 1
Input 4k video details.

Input # (Input No.)	Frame Width	Frame Height	Frame rate [fps]	Length [s]	Size [MB]	Bitrate [kbps]
1	3840	2160	23.98	24	62.9	22003
2	3840	2160	23.98	63	165	21722
3	4096	2160	23.98	26	58.4	18165
4	3840	2160	23.98	76	200	21808
5	3840	2160	23.98	100	262	21814
6	3840	2160	23.98	43	114	21848

Besides aomenc-AV1, rav1e-AV1 is selected as one of the new AV1 formats for experimental analysis and comparison with aomenc-AV1. For the purpose of experiment NotEnoughAV1Encodes tools is applied, as well as ffmpeg [8, 14]. Only one-pass coding is applied here.

Performace evaluation is performed by calculation of three most common metrics for reference-based video quality assessment. The first one is PSNR (Peak Signal-to-Noise Ratio) representing the ratio between the maximum signal power and power of noise estimation. It can be calculated using MSE (Mean Squared Error), having in mind the maximum pixel value, *MAX*:

$$MSE = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[I_q(i,j) - I(i,j)\right]^2,$$
$$PSNR = 20\cdot\log_{10}\left(MAX / \sqrt{MSE}\right). \tag{1}$$

In (1), frame I of size $M\times N$ belongs to the input, while I_q is a reencoded version. The PSNR is given in decibels (dB), where higher value means higher coded video quality.

Besides the most common PSNR metric, SSIM or Structural Similarity is also calculated. This index is a full-reference metric showing how similar the corresponding frames are. In comparison to PSNR and MSE, it is not based on calculating the absolute error, but considers structural similarity based on perceptual approach. If two image parts of the same size, x and y, are compared, SSIM (Structural Similarity Index) can be measured as:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)}{(\mu_x^2 + \mu_y^2 + c_1)}\cdot\frac{(2\sigma_{xy} + c_2)}{(\sigma_x^2 + \sigma_y^2 + c_2)}, \tag{2}$$

where μ_x and μ_y are averages of x and y, respectively, σ_x^2 and σ_y^2 are variances of x and y, respectively, σ_{xy} covariance and c_1 and c_2 are chosen parameters

[20]. SSIM is presented in accordance to scale from 0 to 100. Higher SSIM means that higher structural resemblance exist, which can be related to higher quality.

One of the popular approaches for video assessment is VMAF (Video Multi-Method Assessment Fusion). It is a perceptual video quality assessment project for obtaining a better practical perceptual quality metric for video [21]. Also, it is a Netflix on-going project for improved evaluation based on other metrics, like visual information fidelity, detail loss metric, mean co-located pixel difference and anti-noise signal-to-noise ratio. VMAF value belongs to range from 0 to 100. Here, PSNR, SSIM and VMAF are metrics calculated using Python framework and ffmpeg-quality-metrics package [10, 22].

Metrics are calculated per frame and for each component (Y, U, V), where of interest here is mainly average values for luminent (Y) and two chrominanse components (U and V), as well as values specifically for Y component [10]. Thus, PSNR and SSIM are calculated as average (AVG) values for all components (PSNR AVG, SSIM AVG), as well as for Y component (PSNR Y, SSIM Y). PSNR, SSIM and VMAF are calculated for six 4k sequences.

The experimental analysis is consisted of three phases. In the first phase, the measurements are made only for aomenc-AV1 using the abovementioned metrics, PSNR, SSIM and VMAF. The second phase is performed in order to compare metric values obtained for aomenc-AV1 and rav1e-AV1 per frame, in the case of fixed Q parameter choice. In the third phase, the comparison is made between aomenc and rav1e implementation for seven different Q factors. The aim is to compare a reference-like approach (aomenc) with a general usage tool (rav1e) for obtaining AV1 format, and to observe possible differences for 4k video.

4 Experimental Results

In this paper, analysis of the 4k video traffic is performed using AV1 software implementations called aomenc and rav1e. Three metrics are used for the purpose of comparison between the coded video sequences and the reference H.264 encoded video sequence. The output AV1 sequence in each experiment has the extension .mkv (Matroska Media Container).

4.1 Quality assessment results for aomenc coding

In the first phase, a 4k video sequence is observed for the case when aomenc-AV1 coding is applied. Using the reference-like tool, the coding of original mp4 video sequence is performed using seven constant quality factors (Q). For such coded sequences, the metrics are calculated by comparison with the source (original) sequence. In Fig. 3, the results for Video No. 1 are shown where different constant quality factors are applied. The results represent the calculated PSNR and SSIM metric values.

An increase in the value of Q corresponds to a decrease in the quality of the video sequence. PSNR AVG values represent the mean value of PSNR obtained for frames, averaged for all components (Y, U, V), while PSNR Y represents the average value of the luminescent component (Y) video frames. Values are expressed in decibels (dB). It can be noticed that the PSNR AVG and the PSNR Y are monotonically decreasing functions of Q, and that their difference is an approximately constant function. The difference is not largely affected by the change of the constant quality factor. The PSNR AVG values are expected to be higher because they, in addition to the luminescent Y, include chrominance components U and V.

On the other hand, in Fig. 3 corresponding values of the SSIM AVG and the SSIM Y are also shown. Similarly, as in the case of PSNR, SSIM AVG represents the mean value of SSIM obtained for all frames and components (Y, U, V). The average value of the luminescent component (Y) video frames is noted as SSIM Y. Both SSIM AVG and SSIM Y are expressed as percentages (%), for various constant quality factors. The SSIM values decrease with increasing constant quality factor, while the difference between these two values (SSIM AVG and SSIM Y) is significantly larger at higher Q values.

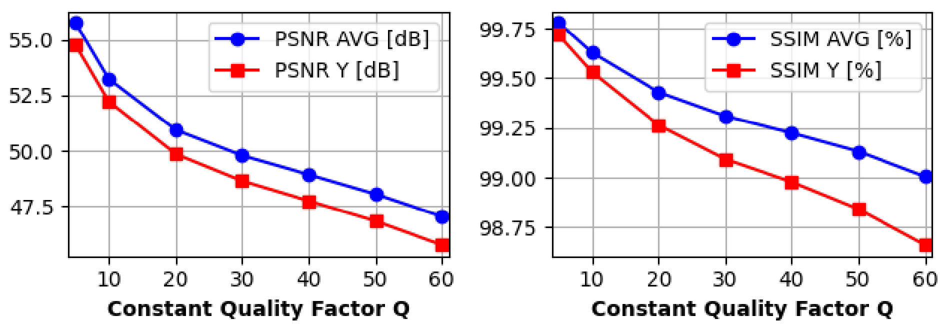

Fig. 3 – *PSNR AVG (average PSNR) and PSNR Y (PSNR for Y- luminant component) obtained for different quality (Q) factors (left); SSIM AVG (average SSIM) and SSIM Y (SSIM for Y- luminant component) obtained for different quality (Q) factors (right).*

The calculated VMAF metric averaged on the frame level for each constant quality parameter Q is presented in Fig. 4. The VMAF values are shown in Fig. 4 for coded Video No. 1. The VMAF represents a decreasing function of constant quality parameter. The decrease in VMAF values is larger at higher Q values, i.e. when there is a drastic reduction in the quality of coded video sequence.

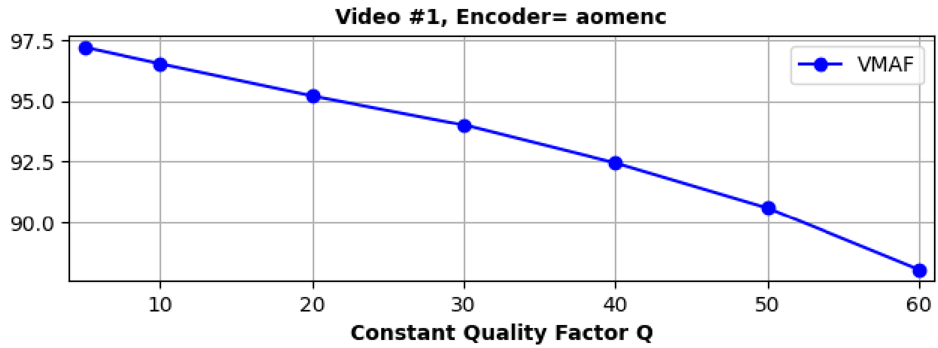

Fig. 4 – *VMAF values obtained for coded Video No.* 1 *using aomenc and seven Q factors.*

4.2 Comparison results between aomenc and rav1e coding for a fixed quality choice

In the second phase, video sequences after coding with two different implementations of AV1 are compared. Namely, these implementations represent the aomenc-AV1 and the rav1e-AV1 encoder. In contrast to the previous phase, when the mean values of metrics are calculated at the frame level, in this phase time series corresponding to the metric values PSNR AVG and SSIM AVG are obtained. Both PSNR AVG and SSIM AVG are calculated for every frame. In Fig. 5 they are obtained for Video No. 1 as average values for the components (Y, U, V). Here, the quality parameter Q is set to a fixed value 30 (Q = 30).

It is evident that the higher quality is on the side of rav1e-AV1 implementation for PSNR AVG evaluation. The mean values per frame for the cases of aomenc-AV1 and rav1e-AV1 are 49.79 dB and 53.64 dB, respectively. The standard deviations for these cases are 0.74 dB and 1 dB, respectively. Higher mean value is obtained for rav1e-AV1 implementation, while standard deviations are similar.

The mean values of SSIM (SSIM AVG) per frame for aomenc-AV1 and rav1e-AV1 are 99.31 % and 99.64 %, respectively. The standard deviations for these cases are 0.19 % and 0.1 %, respectively. By observing the calculated SSIM metrics, it can be pointed out that the higher mean value and the lower dispersion correspond to the rav1e-AV1 implementation. Moreover, it can be noticed that in the first sixty frames there are points where significantly lower SSIM values occur in the case of aomenc-AV1. Only, one-pass coding is performed here. The difference in the remaining frames is smaller. Also, it is shown that there are cases where successive frames have same SSIM AVG values. This can be interpreted as saturation occurrence for these values.

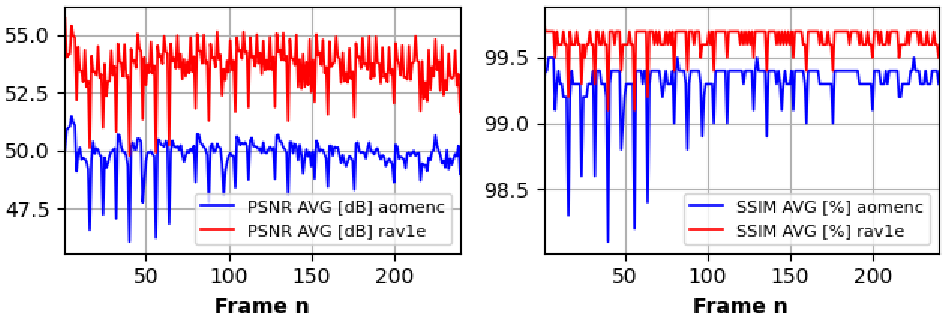

Fig. 5 – *Average PSNR (PSNR AVG) for aomenc and rav1e codec (left) for Q* = 30; *average SSIM (SSIM AVG) for aomenc and rav1e codec (right) for Q* = 30.

VMAF metrics for video sequence Video No. 1 is shown is Fig. 6. The values are presented for the case of aomenc and rav1e, and constant quality factor Q = 30.

The mean values of VMAF per frame in the case of aomenc-AV1 and rav1e-AV1 are 94.02 and 96.54, respectively. The calculated standard deviations for these cases are 0.62 and 0.44, respectively. The mean value for the rav1e-AV1 implementation is higher and the standard deviation is lower, showing better performance of rav1e in comparison to aomenc. This can be explained by the fact that the rav1e-AV1 implementation yields a video sequence that is closer to the original/source sequence quality, indicating generally better results.

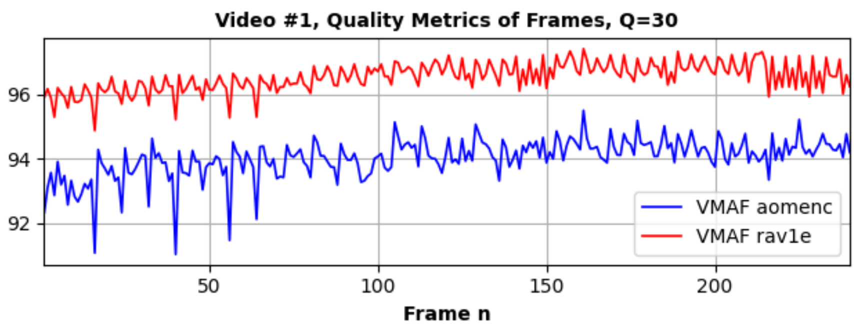

Fig. 6 – *VMAF based comparison between aomenc and rav1e for Q* = 30.

4.3 Comparison results between aomenc and rav1e for different constant quality factors

In the third phase, the comparison is made between aomenc-AV1 and rav1e-AV1, where seven values of the constant quality parameter are used. This is presented by PSNR AVG and SSIM AVG results in Fig. 7, where Video No. 1

represents the source video. These values are averaged by components and frames. Higher values are obtained when rav1e-AV1 is applied. PSNR based differences between rav1e-AV1 and aomenc-AV1 can be considered approximately constant as constant quality factor changes. In Fig. 7, SSIM averaged values are also shown. Higher values of SSIM AVG are achieved when the rav1e-AV1 is implemented. The difference with respect to aomenc-AV1 increases while constant quality factor Q increases.

VMAF values in Fig. 6 are calculated per frames for a fixed constant quality factor. Higher VMAF values are obtained in the case of rav1e. Also, smaller standard deviation is calculated. In Fig. 8, VMAF is shown for seven different Q values and the two implementations.

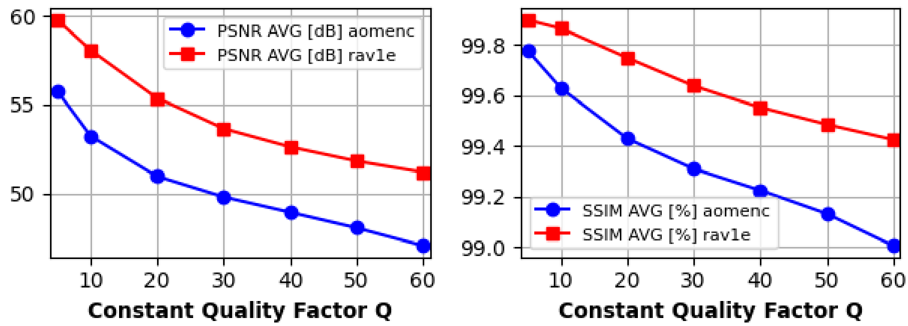

Fig. 7– *Average PSNR for aomenc and rav1e codec (left); Average SSIM for aomenc and rav1e codec (right) obtained for different Q factors.*

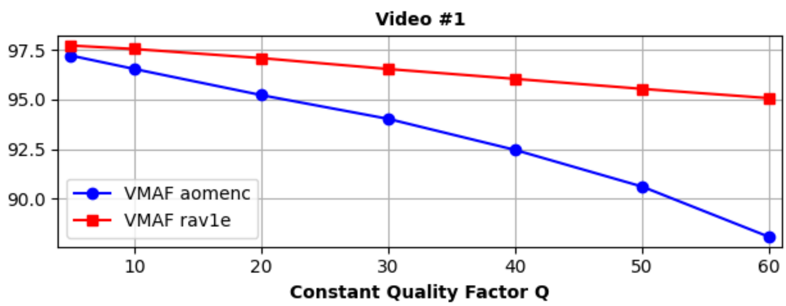

Fig. 8– *Average VMAF for aomenc and rav1e codec obtained for different Q factors.*

VMAF values in Fig. 8 are averaged, where the rav1e implementation gives better results. Moreover, it can be observed that for rav1e these values decreases much slower for increasing Q. Moreover, VMAF metric compared to PSNR and SSIM shows evident trend of the difference changes for increasing Q factor.

Finally, in order to avoid averaged versions of VMAF for better insight into obtained variations, the calculation is performed without usual averaging for both implementations. Obtained statistics for each constant quality factor Q can be observed for different source video content in Fig. 9. Namely, by using box plot diagrams several measures are shown at the same time: median, mean, maximum, minimum, as well as boundaries representing the first and the third quartile intervals.

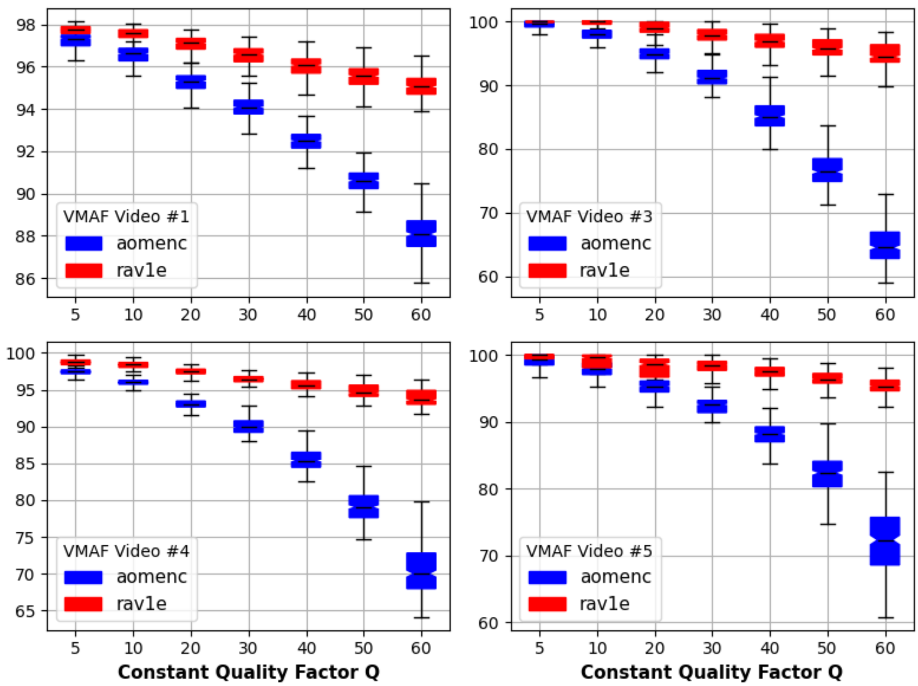

Fig. 9 – *Box plot representations of VMAF values for aomenc-AV1 and rav1e-AV1 encoders, obtained for seven Q factors and different source sequences.*

It can be seen that similar VMAF behavior of AV1 implementations has been shown for different 4k source video content. This means that there is a slower decrease of VMAF with increasing Q. In other words, for lower quality of coded video represented by higher Q, better results are obtained for rav1e compared to the reference aomenc. More intense decrease exists for the aomenc choice with increasing constant quality factor. Also, there is an increase of interquartile difference for aomenc-AV1 for higher Q factor values.

Difference between mean values of PSNR AVG per frame (noted as $\overline{PSNR\,AVG}$) for the two implementations are calculated, as well as difference between mean values of SSIM AVG per frame (noted as $\overline{SSIM\ AVG}$) for rav1e and aomenc. These PSNR and SSIM based differences,

$$\overline{PSNR\,AVG}_{rav1e} - \overline{PSNR\,AVG}_{aomenc},$$

$$\overline{SSIM\ AVG}_{rav1e} - \overline{SSIM\ AVG}_{aomenc},$$

are calculated for each Q, and for six source video sequences. The trends approximated as linear models are shown in Fig. 10. It can be seen that for Video No. 1 PSNR based difference is almost constant. Generally, there is an increase of the differences from both PSNR and SSIM aspect with increasing factor Q. In the case of PSNR, all differences are positive. Nevertheless, if a source video is full of details and different textures as in Video No. 2, a decreasing effect may occur in the SSIM based difference. For all tested video material, except Video No. 2, positive SSIM based difference can be seen in Fig. 10 for Q factor higher than 30.

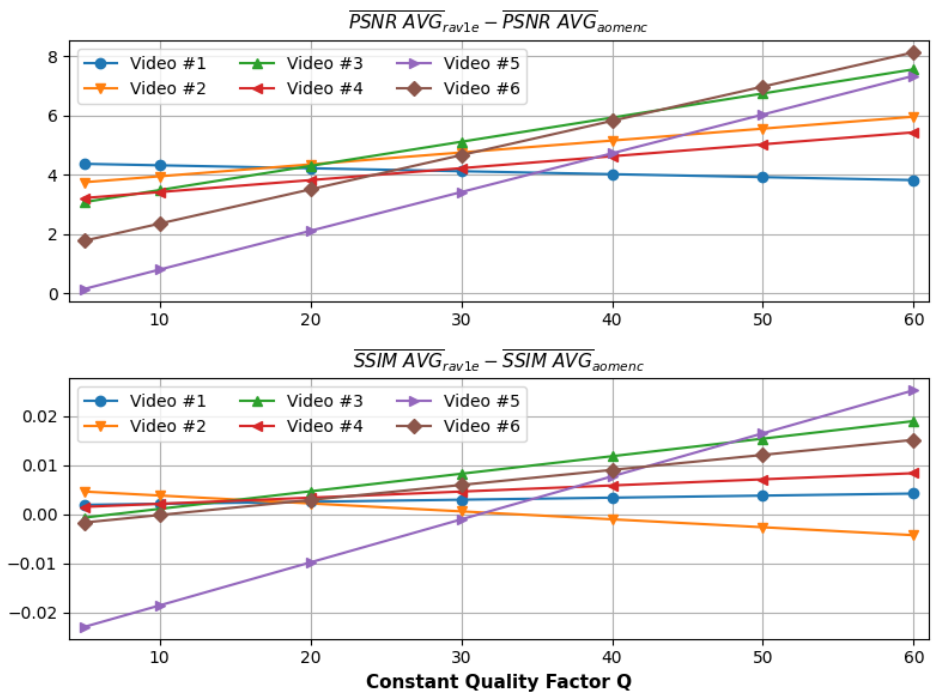

Fig. 10 – *Trends of PSNR AVG (upper) and trends of SSIM AVG differences (lower) for seven Q factors and different source sequences.*

5 Conclusion

This paper presents experimental analysis based on new AV1 (Alliance for Open Media Video 1) format for 4k video content. Video sequences were analyzed using two available AV1 implementations, aomenc and rav1e. Constant quality factor was varied in order to test these implementations, where three different metrics were implemented: PSNR, SSIM and VMAF. The results in this paper showed that higher quality (higher similarity to the source) for different factor values is obtained for rav1e compared to the aomenc reference. The improvements made so far in AV1 format are evident.

In the future research, it is necessary to analyze: the influence on coding results for different bitrate constraints, various CPU (Central Processing Unit) speed settings, multipass coding effects, as well as video coding results for different bit depth choices. Moreover, due to a high number of AV1 formats, this should be analyzed for various implementations and video content.

6 Acknowledgment

This research is supported by Serbian Ministry of Education, Science and Technological Development.

7 References

[1] CISCO Visual Networking Index: Forecast and Methodology, 2017–2022, CISCO, November 2018.

[2] I. Trow: AV1: Implementation, Performance, and Application, SMPTE Motion Imaging Journal, Vol. 129, No. 1, January 2020, pp.51–56.

[3] Alliance for Open Media, http://aomedia.org/, (last accessed 25.04.2021.)

[4] J. Shulman: Bitmovin's Video Top Video Technology Trends 2020, Bitmovin, March 2020. https://bitmovin.com/video-technology-trends/ (last accessed 15.07.2020.)

[5] J. Ozer: The Streaming Codec Landscape in 2021, Wowza Media Systems, January 2021, https://www.wowza.com/blog/streaming-codec-landscape-2021 (last accessed 20.04.2021.)

[6] F. Zhang, A.V. Katsenou, M. Afonso, G. Dimitrov, D.R. Bull: Comparing VVC, HEVC and AV1 Using Objective and Subjective Assessments, March 2020, pp. 1–9.

[7] Y. Chen, D. Mukherjee, J. Han, A. Grange, Y. Xu, S. Parker, C. Chen, H. Su, U. Joshi, C.- H. Chiang, Y. Wang, P. Wilkins, J. Bankoski, L. Trudeau, N. Egge, J.- M. Valin, T. Davies, S. Midtskogen, A. Norkin, P. de Rivaz, Z. Liu: An Overview of Coding Tools in AV1: The First Video Codec from the Alliance for Open Media, APSIPA Transactions on Signal and Information Processing, Vol. 9, February 2020, pp. 1–15.

[8] M. Xu, B. Jeon: User-Priority Based AV1 Coding Tool Selection, IEEE Transactions on Broadcasting, 2021, pp. 1-10.

[9] J. Han, B. Li, D. Mukherjee, C.- H. Chiang, A. Grange, C. Chen, H. Su, S. Parker, S. Deng, U. Joshi, Y. Chen, Y. Wang, P. Wilkins, Y. Xu, J. Bankoski: A Technical Overview of AV1, Proceedings of the IEEE, February 2021, pp. 1–25.

[10] FFmpeg, https://ffmpeg.org/ (last accessed 25.04.2021.)

[11] FFmpeg - Libaom, https://trac.ffmpeg.org/wiki/Encode/AV1 (last accessed 04.04.2021.)

[12] M. Milivojevic, D. Dujkovic, A. Gavrovska: Constant Quality Mode 4k Video Comparison Using AV1 Reference Tool, Proceedings of the International Conference on Electrical, Electronic and Computing Engineering, Belgrade, Serbia, September 2020, pp. 336–339.

[13] Rav1e, https://github.com/xiph/rav1e, (last accessed 04.04.2021.)

[14] NotEnoughAV1Encodes,
https://github.com/Alkl58/NotEnoughAV1Encodes, (last accessed 04.04.2021.)

[15] AV1 OpenBenchmarking,
https://openbenchmarking.org/suite/pts/av1, (last accessed 25.04.2021.)

[16] D.R. Marković, A.M. Gavrovska, I.S. Reljin: 4K Video Traffic Prediction Using Seasonal Autoregressive Modeling, Telfor Journal, Vol. 9, No. 1, 2017, pp. 8–13.

[17] A.M. Gavrovska, M.S. Milivojevic, G. Zajic, I.S. Reljin: Video Traffic Variability in H.265/HEVC Video Encoded Sequences, Proceedings of the 13th Symposium on Neural Networks and Applications (NEUREL), Belgrade, Serbia, November 2016, pp. 109–112.

[18] A.M. Gavrovska, M.S. Milivojevic, G. Zajic: Analysis of SVT-AV1 Format for 4k Video Delivery, Proceedings of the 28th Telecommunications Forum (TELFOR), Belgrade, Serbia, November 2020, pp. 1–4.

[19] Pexels, https://www.pexels.com, (last accessed 12.04.2021.)

[20] A. Gavrovska, A. Samovic: Evaluation of Image Dynamic Range Reduction, PosTel 2017, Simposium Belgrade, Serbia, December 2017, pp. 221–230.

[21] VMAF - Video Multi-Method Assessment Fusion,
https://github.com/Netflix/vmaf, Netflix, (last accessed 25.04.2021.)

[22] FFmpeg Quality metrics,
https://pypi.org/project/ffmpeg-quality-metrics/, (last accessed 04.04.2021.)