Article

Methodology for Modeling and Comparing Video Codecs: HEVC, EVC, and VVC

Stefano Battista[ID]**, Massimo Conti**[ID] **and Simone Orcioni ***[ID]

Department of Information Engineering, Università Politecnica delle Marche, 60131 Ancona, Italy;
s.battista@univpm.it (S.B.); m.conti@univpm.it (M.C.)
* Correspondence: s.orcioni@univpm.it; Tel.: +39-071-2204543

check for updates

Abstract: Online videos are the major source of internet traffic, and are about to become the largest majority. Increasing effort is aimed to developing more efficient video codecs. In order to compare existing and novel video codecs, this paper presents a simple but effective methodology to model their performance in terms of Rate Distortion (RD). A linear RD model in the dB variables, Peak Signal-to-Noise Ratio (PSNR) and Bitrate (BR), easily allows us to estimate the difference in PSNR or BR between two sets of encoding conditions. Six sequences from the MPEG test set with the same resolution, encoded at different BR and different Quantization Parameters, were used to create the data set to estimate each RD model. Three codecs (HEVC, EVC, and VVC) were compared with this methodology, after estimating their models. Fitting properties of each model and a performance comparison between the models are finally shown and discussed.

Keywords: video coding; video codecs; video compression; codec modeling and comparison

1. Introduction

Since the advent of digital video technology in the late 1980s, the amount of video data delivered over the global communication network has been increasing constantly. Recent data show that video accounted for 75% of the Internet Protocol traffic in 2017 and it is forecast to exceed 80% in 2022 [1].

Video compression schemes are thus essential for significantly reducing the data capacity for storage, or the data bandwidth for transmission of such big data. Those are the motivations behind the over 30 years of development of video coding standards in ITU (since H.261 in 1988) and in ISO (since MPEG-1 in 1993).

The current state-of-the-art video coding standards in the context of ITU, ISO and industrial fora are: High Efficiency Video Coding (HEVC/H.265), Essential Video Coding (EVC), Versatile Video Coding (VVC/H.266), and AV1. HEVC/H.265 is the joint ISO/ITU video coding standard finalized in 04/2013 [2]. EVC is the ISO video coding standard at the stage of Final Draft International Standard and scheduled for finalization in 2020 [3]. VVC/H.266 is the joint ISO/ITU video coding standard at the stage of Final Draft International Standard and is scheduled for finalization in 2020 [4]. AOM AV1 is an open specification from the industrial forum Alliance for Open Media.

Ohm et al. [5] published in 2012 a study of the state-of-the-art video codecs, comparing MPEG-2/H.262, H.263, MPEG-4, AVC/H.264 with the most recent standard HEVC/H.265 that was published in 2013. The main outcomes were the usage of 50% less bitrate of HEVC with respect to AVC for the same Peak Signal to Noise Ratio (PSNR), and an even higher subjective quality improvement.

Similarly, Dissanayake and Abeyrathna [6] in 2015 compared the bitrate/quality performance of AVC and HEVC in a broadcast environment, taking into account also the increase in processing complexity to achieve the 50% saving in bitrate of HEVC.

Gros et al. [7] in 2013 extended the comparison to AVC, HEVC and the proprietary codec VP9 published in 2013, the same year as HEVC. The paper reported bitrate savings for HEVC of 39% over AVC and 43% over VP9. The paper also reported a coding efficiency for VP9 inferior to both AVC by 8.4% and HEVC by 79.4% in terms of average bitrate overhead at the same objective quality, with a processing complexity greater by a factor 100 when comparing VP9 and x264 (open source AVC), and lower by a factor 7 when comparing VP9 and HM (reference HEVC).

Barman and Martini [8] in 2017 presented an objective evaluation of the eight most popular games encoded using AVC, HEVC, and VP9 encoders for live game video streaming applications. The results are reported in terms of three objective video quality metrics (PSNR, SSIM, VIFp), Bjontegaard-Delta Bit-Rate (BD-BR) analysis, and encoding time. HEVC provided the best compression efficiency in terms of BD-BR analysis, with an encoding time being three times slower than AVC, and AVC provided better compression than VP9 with an encoding time four times faster than VP9.

Gros et al. [9] in 2018 further extended the comparison to HEVC, JEM (preliminary model for VVC), VP9 (proprietary), and AV1 (AOM open evolution of VP9). The authors obtained the following results: AV1 achieved an average bitrate savings of 17% relative to VP9 at the cost of a factor 117 in encoder run time. JEM (model for VVC) achieved an average bitrate saving of 30% relative to HEVC at the cost of a factor of 11 in encoder run time. AV1 produced an average bitrate overhead of more than 100% relative to JEM at the same objective reconstruction quality besides a factor of three in encoder run time. Even in a two-pass rate-control mode, AV1 had an overhead of 55% relative to JEM (VVC) and 10% relative to HEVC.

In a different direction, Katsavounidis and Guo [10] in 2018 presented a new methodology for an objective comparison of video codecs. Using Video Multimethod Assessment Fusion (VMAF), an open-source perceptual video quality metric, the paper proposes a visual perceptual optimization of any video codec in terms of PSNR and VMAF. The methodology is applied to encoder implementations of AVC, HEVC, and VP9. The paper reports advantages and disadvantages of different encoders for different bitrate/quality ranges and for a variety of contents.

This work presents a methodology for modeling the performances of video codecs. The modeling and comparison are based on the estimation of the Rate Distortion (RD) curves for each of the video codecs using experimental data from a set of video sequences at the same spatial resolution, in terms of Bitrate (BR) and PSNR. The methodology was applied in order to compare the performances of three different video codecs: MPEG-H HEVC/H.265, MPEG-5 EVC, MPEG-I VVC/H.266, by using six video sequences at Ultra HD (UHD) resolution. The modeling and comparison of these state-of-the-art video codecs are achieved by first studying the behavior of the codecs in terms of BR and PSNR variations depending on the Quantization Parameter (QP) values and under similar encoding configurations, and then studying an averaged model for each video codec for a set of sequences at the same spatial resolution. Using such modeling of different video codecs in terms of RD curves, it is possible to compare their performances in terms of Bitrate difference at the same PSNR, or PSNR difference at the same BR.

The main novelty and advantage of the proposed modeling and comparison methodology consists of the possibility to compare, both numerically and graphically, the trend of the RD curves in a range of BR and PSNR. The widely adopted BD-BR algorithm [11,12] provides a single numerical value for each sequence on a given range of Bitrates, and a single numerical value for the average of the RD performance among several sequences. The proposed method, instead, provides an RD model for each sequence and an averaged RD model, derived from the set of sequence models: this allows a more significant analysis of the RD characteristics for each codec at different bitrates.

The same methodology can also be extended to modeling and comparing the codecs performance with respect to other metrics, such as VMAF or Mean opinion score (MOS).

The paper is organized as follows. Section 2 explains how the data set, used for the estimation of the codec RD models, was obtained from a chosen set of sequences. Section 3 introduces, in Section 3.1 the methodology for modeling the RD behavior of the codecs, for each sequence, and the realization of

the averaged RD codec model. In Section 3.2 the results of model estimation are presented. Section 4 explains how to use the averaged codec model to compare the codecs in terms of PSNR and Bitrate. A Comparison with the well established BD-BR algorithm is given in Section 5, and eventually conclusions are drawn in Section 6.

2. Experimental Data Set for the Codec Modeling

With the aim of introducing a methodology for the comparison of the performances of video codecs, this Section presents the procedure followed to provide the data set used for the estimation of the codec models. Six sequences with different frame rates (50 and 60 frames per second) and the same resolution (UHD, 3840×2160) were chosen for producing the data set. Such sequences are a subset from the MPEG test set:

DaylightRoad at 3840×2160 and 60 fps;
CatRobot   at 3840×2160 and 60 fps;
Fortnite   at 3840×2160 and 60 fps;
ParkRunning at 3840×2160 and 50 fps;
FlyingBirds   at 3840×2160 and 60 fps;
SunsetBeach   at 3840×2160 and 60 fps.

As an example of an application of the methodology, the following video codecs were considered:

- HEVC, from MPEG, as implemented in the Test Model HM 19;
- EVC, from MPEG, as implemented in the Test Model ETM 6;
- VVC, from MPEG, as implemented in the Test Model VTM 9.

For the comparison of the three codecs, the PSNR as a function of BR has been chosen as RD curve. The main criterion is to operate at fixed QP (i.e., encoding the whole sequence with the same QP) from a set of four predefined QP values, applied to each sequence with each codec: 27, 32, 37, 42. Of course, the resulting data can be extended to more sequences at the same resolution. For similar comparative analysis see [9,13,14]. By using such QP values the corresponding PSNR_{YUV} values were calculated according to the following expression

$$\text{PSNR}_{\text{YUV}} = \left(6\,\text{PSNR}_{\text{Y}} + \text{PSNR}_{\text{U}} + \text{PSNR}_{\text{V}}\right)/8\,. \tag{1}$$

In the following Figures, to avoid subscripts in the labels, the PSNR_{YUV} will be indicated as YUV-PSNR. Each PSNR in the right part of (1) is calculated as:

$$\text{PSNR} = 10\log\frac{\left(2^{\text{B}} - 1\right)^2}{\text{MSE}}\,, \tag{2}$$

where, for a picture size of M rows by N columns of pixels,

$$\text{MSE} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\text{original}(i,j) - \text{coded}(i,j)\right)^2\,, \tag{3}$$

where B is the number of bits per sample of Luminance (Y) and Chrominance (U, V), and $\text{original}(i,j)$ and $\text{coded}(i,j)$ are the Y, U and V values of the pixel at position (i,j) in the original and coded pictures, respectively.

As an example, Table 1 reports the data resulting from encoding the sequence DaylightRoad. The QP values and respective BR and PSNR_{YUV} values are shown in the Table for the encodings with the three codecs: HEVC, EVC, and VVC. From first inspection, we can see that the experimental data in Table 1 show different characteristics. For the same range of QP's, the bitrates obtained are higher for HEVC with respect to EVC, and for EVC with respect to VVC. However, for the same QP, the PSNR values are almost always increasing, going from HEVC to VVC, with a maximum difference of 0.48 dB

(QP = 42) and decreasing, going from HEVC to EVC, with a maximum difference of 0.16 dB (QP = 32). Furthermore, we can expect these results to change by using a different sequence.

Table 1. DaylightRoad sequence.

	HEVC		EVC		VVC	
QP	**BR**	**PSNR$_{\text{YUV}}$**	**BR**	**PSNR$_{\text{YUV}}$**	**BR**	**PSNR$_{\text{YUV}}$**
	(Mbps)	**(dB)**	**(Mbps)**	**(dB)**	**(Mbps)**	**(dB)**
27	18.932	36.52	12.794	36.41	12.367	36.44
32	9.721	35.86	6.557	35.70	6.718	35.89
37	4.993	34.91	3.288	34.76	3.559	35.12
42	2.800	33.66	1.937	33.63	2.141	34.14

These observations show the need for a consistent and repeatable methodology for a fair comparison of such data sets in a given range of BR and PSNR. The proposed methodology is described in the following section.

3. Codec Modeling: PSNR VS. Bitrate

After collecting the data set for the different sequences encoded using different codecs, in the present Section the methodology for modeling the RD behavior of the codecs for the chosen set of sequences is described.

3.1. Sequence Model and Averaged Codec Model

With the four couples of (PSNR$_{\text{YUV}}$, BR), obtained in the previous Section for each sequence, it was possbile to estimate the coefficients of a PSNR model as a function of BR. Since the RD curve of PSNR$_{\text{YUV}}$ as a function of BR shows a logarithmic trend, according to [11,12], in order to work with linear models, the logarithmic variable BR$_{\text{dB}}$ derived from BR is defined as follows:

$$BR_{dB} = 10 \log(BR/BR_0) \,, \tag{4}$$

where BR$_0$ is a normalization constant, defined as BR$_0$ = 1 bps, in order to have a dimensionless logarithmic argument.

The proposed model PSNR$_{\text{YUV}} = f(BR_{dB})$ is then:

$$PSNR_{YUV} = A + B \, BR_{dB} \,, \tag{5}$$

that is a linear function in the dB bitrate. The inverse model, BR$_{dB} = g(PSNR_{YUV})$, results as follows

$$BR_{dB} = C + D \, PSNR_{YUV} \,, \tag{6}$$

with $C = -A/B$, and $D = 1/B$.

Using the models described above, it is possible to compute an averaged model [15,16] for each set of sequences, e.g., one single model for all sequences at 3840×2160 resolution. Such an averaged model will be more significant for increasing the number of input sequences used to compute the experimental data and the related sequence models for PSNR$_{\text{YUV}}$ as a function of BR$_{\text{dB}}$.

The motivation for using a simple average of least-square models is that with linear models, the average of the model parameters is perfectly coincident with the optimal least-square model of the aggregate data from the different series. In other words, if the linear model (5) is computed as the average of a set of linear models across a number of data sets, exactly the same parameters of the linear model computed over the union of the same data sets are obtained.

3.2. Model Estimation Results

This section reports the estimation results obtained with the following configurations:

- three codecs: HEVC (HM 16), EVC (ETM 6), VVC (VTM 9);
- six sequences: DaylightRoad, CatRobot, Fortnite, ParkRunning, FlyingBirds, SunsetBeach;
- four QP values: 27, 32, 37, 42.

From the data points reported in Table 1, and from equivalent data for the other test sequences, the parameters *A* and *B* of the linear RD model (5) were computed for each sequence and for each codec.

To estimate the general behavior of a specific video codec, HEVC in this first case, the technique of model averaging was used, i.e., computing the average of the parameters of the linear models across sequences of the same class of resolution. The coefficients of the sequence RD models and those resulting from model averaging for the HEVC codec are reported in Table 2. Figure 1 shows data points for the 3840×2160 sequences and the averaged RD model.

Table 2. Models.

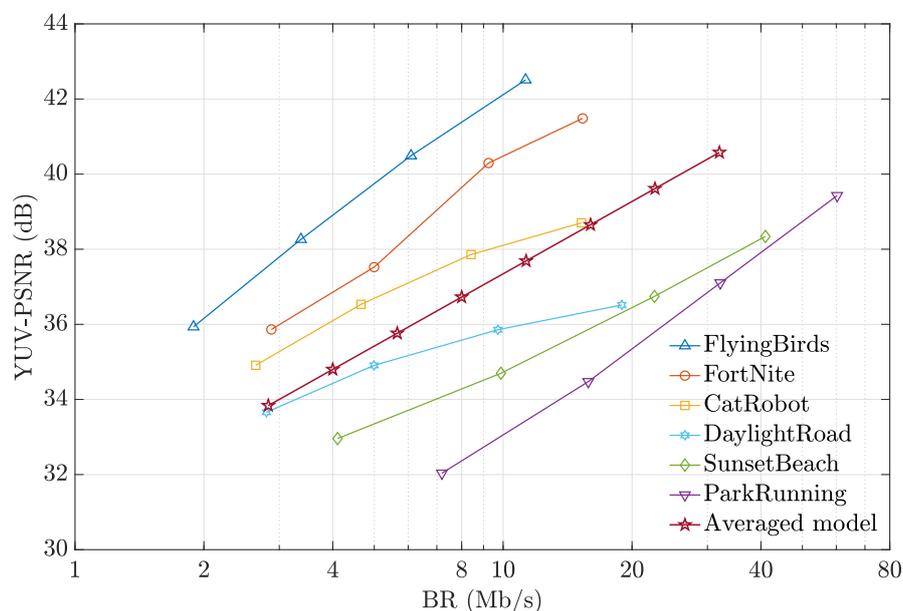| Sequence	HEVC		EVC		VVC	
	A	*B*	*A*	*B*	*A*	*B*
FlyingBirds	−17.26	+0.8491	−14.38	+0.8197	−15.43	+0.8368
FortNite	−16.17	+0.8051	−12.32	+0.7657	−11.45	+0.7593
CatRobot	+3.034	+0.4996	+4.189	+0.4929	+8.087	+0.4378
DaylightRoad	+11.89	+0.3406	+12.79	+0.3344	+15.41	+0.2983
SunsetBeach	−2.848	+0.5394	−0.9792	+0.5184	−0.6346	+0.5140
ParkRunning	−23.24	+0.8043	−17.89	+0.7378	−18.70	+0.7472
Average	−7.432	+0.6370	−4.767	+0.6115	−3.786	+0.5989

Figure 1. High Efficiency Video Coding (HEVC): PSNR–BR data and linear RD models for 3840×2160 resolution sequences. The averaged RD model is also displayed as red line with star markers.

As above, the linear RD model (5) for each sequence, and the average for each class of sequences with EVC coding were computed. The coefficients of the linear RD models and of the averaged RD model for the EVC codec are also reported in Table 2. Figure 2 shows the data and the estimated averaged RD model relative to the EVC codec, for the 3840×2160 sequences.
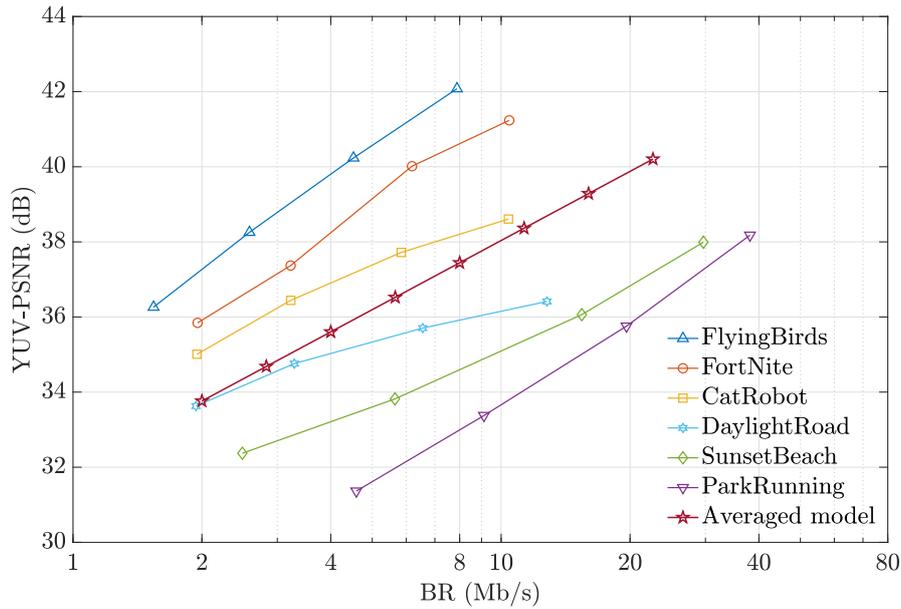
Figure 2. Essential Video Coding (EVC): Peak Signal-to-Noise Ratio (PSNR)–Bitrate (BR) data and linear Rate Distortion (RD) models for 3840×2160 resolution sequences. The averaged RD model is also displayed as red line with star markers.

The coefficients of the linear RD models for the VVC codec are reported in the last two columns of Table 2. Figure 3 shows data relative to the sequences and the estimated averaged model.
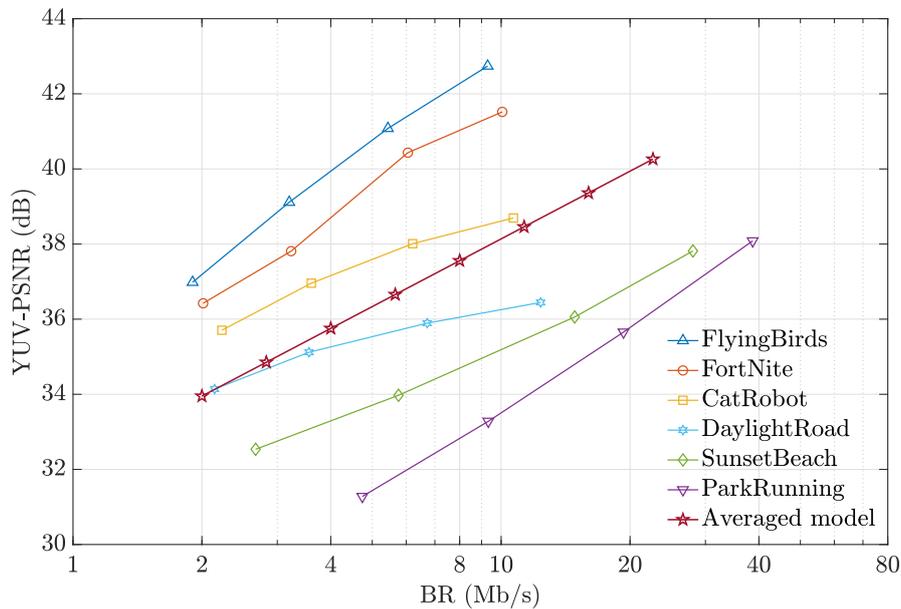

Figure 3. Versatile Video Coding (VVC): PSNR–BR data and linear RD models for 3840×2160 resolution sequences. The averaged RD model is also displayed as red line with star markers.

As can be seen from Table 2, the angular coefficients (B) of the linear model of each sequence are quite similar among the different codecs, while the intercepts with the ordinate axis at zero BR_{dB} (A) have greater variation among the codecs. If the angular coefficients were equal, the intercept A would determine which model has the higher PSNR, however, since the A coefficient is the model value at zero BR_{dB}, i.e., for BR equal to 1 bps, the angular coefficient B must be multiplied by 70 to obtain the model values at 10 Mbps, i.e., around the middle of the range of BR values; accordingly also small differences in B matter.

The procedure described above to estimate the parameters of the model $\text{PSNR}_{\text{YUV}} = f(\text{BR}_{\text{dB}})$ shows a good match of the approximating functions to the experimental data, with a coefficient of determination (R^2) between the actual data and the theoretical model between $R^2 = 0.934024$ and $R^2 = 0.999708$ for all sequences.

Comparing the behavior of the codecs under the same test conditions, using the procedure and the modeling described above, we can obtain a realistic estimate of the performances of each codec in terms of:

- QP values used for coding;
- Bitrate resulting from the selected QP;
- PSNR resulting from the selected QP.

4. Comparison of the Codecs in Terms of PSNR and Bitrate

Let us suppose we want to compare the performance of two codecs in a given range of BR values. Identifying the reference codec as "H" and the tested codec as "K", and considering an interval in BR between $\text{BR}_{\text{dB}}^{(1)}$ and $\text{BR}_{\text{dB}}^{(2)}$ (or an interval in PSNR between $\text{PSNR}_{\text{YUV}}^{(1)}$ and $\text{PSNR}_{\text{YUV}}^{(2)}$) the average difference in PSNR (or in BR) between the codecs can be computed with the following procedure.

For the two models, H and K, we apply the linear model (5), and define the difference in PSNR for the extrema of the BR range ($\text{BR}_{\text{dB}}^{(i)}$):

$$
\begin{aligned}
\text{H-PSNR}_{\text{YUV}}^{(i)} &= A_H + B_H\,\text{BR}_{\text{dB}}^{(i)} \\
\text{K-PSNR}_{\text{YUV}}^{(i)} &= A_K + B_K\,\text{BR}_{\text{dB}}^{(i)} \\
\Delta\text{PSNR}_{\text{YUV}}^{(i)} &= \text{H-PSNR}_{\text{YUV}}^{(i)} - \text{K-PSNR}_{\text{YUV}}^{(i)}
\end{aligned}
\tag{7}
$$

Thus, the average difference of PSNR over the given interval of BR results in:

$$
\text{avg}(\Delta\text{PSNR}_{\text{YUV}}) = \frac{\Delta\text{PSNR}_{\text{YUV}}^{(2)} + \Delta\text{PSNR}_{\text{YUV}}^{(1)}}{2} = (A_H - A_K) + (B_H - B_K)(\text{BR}_{\text{dB}}^{(2)} + \text{BR}_{\text{dB}}^{(1)})/2 .
\tag{8}
$$

Conversely, we can also apply the linear model (6), and define the difference in BR for a given value of $\text{PSNR}_{\text{YUV}}^{(i)}$:

$$
\begin{aligned}
\text{H-BR}_{\text{dB}}^{(i)} &= C_H + D_H\,\text{PSNR}_{\text{YUV}}^{(i)} \\
\text{K-BR}_{\text{dB}}^{(i)} &= C_K + D_K\,\text{PSNR}_{\text{YUV}}^{(i)} \\
\Delta\text{BR}_{\text{dB}}^{(i)} &= \text{H-BR}_{\text{dB}}^{(i)} - \text{K-BR}_{\text{dB}}^{(i)}
\end{aligned}
\tag{9}
$$

Thus, the average difference of BR over a given interval of PSNR results in:

$$
\text{avg}(\Delta\text{BR}_{\text{dB}}) = \frac{\Delta\text{BR}_{\text{dB}}^{(2)} + \Delta\text{BR}_{\text{dB}}^{(1)}}{2} = (C_H - C_K) + (D_H - D_K)(\text{PSNR}_{\text{YUV}}^{(2)} + \text{PSNR}_{\text{YUV}}^{(1)})/2 .
\tag{10}
$$

Figure 4 shows graphically the comparison among the different models for 3840×2160 resolution.

From Figure 4 and Table 2 it can be seen that VVC has an average gain of 0.83 dB in PSNR_{YUV} over HEVC in the range of BR between 2 and 32 Mbps. This gain is practically constant over all the range of considered BR values, as can be easily seen from Table 2, that reports very similar coefficients for the two models ($B_{\text{HEVC}} = 0.6370$, $B_{\text{VVC}} = 0.5989$). On the other hand, considering an interval of PSNR_{YUV}, the same result can be seen as an average saving of 25.06% in BR of VVC over HEVC, in the range of PSNR between 30 dB and 46 dB.

Figure 4 and Table 2 also show that EVC has an average gain of 0.72 dB over HEVC, for the same BR range of 2 Mbps to 32 Mbps. Comparing the angular coefficients, we can also note that the gain for EVC is practically constant with respect to HEVC, given that $B_{\text{EVC}} = 0.6115$. All data about the angular coefficients are reported in Table 2.
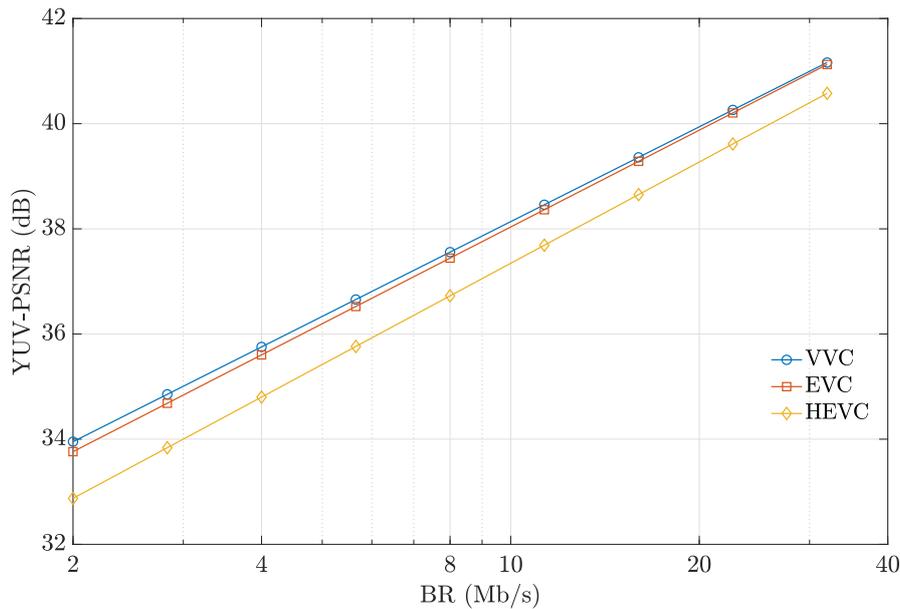
Figure 4. Codec comparison for 3840 × 2160 resolution sequences.

As a summary of the analysis, Table 3 shows the comparisons both in terms of average PSNR difference and average BR percent difference for EVC and VVC, all with respect to HEVC used as a reference.

Table 3. Models comparison with respect to HEVC.

Resolution	EVC	VVC
3840 × 2160	+0.72 dB	+0.83 dB
3840 × 2160	−22.05% BR	−25.06% BR

5. Comparison with the BD-BR Algorithm

This Section compares the results obtained by applying the proposed method with the cited BD-BR algorithm. The relationships (8) and (10) have the same purpose as the BD-BR algorithm [11,12]: estimating differences in PSNR or BR between two corresponding RD curves.

The most common implementation of the BD-BR algorithm to compute the average bitrate savings for a given PSNR range (BD-BR), or conversely the average PSNR gain for a given bitrate range (BD-PSNR), estimates the area between the two RD curves. This is achieved using a piecewise cubic Hermite polynomials (PCHIP) interpolation on the data, with the bitrate measured on a log scale. The area between the two piecewise cubic curves, i.e., the integral between them, is then computed. As an example, the numerical integration can be performed with 1000 equal-sized subintervals, as described in [5]. Table 4 reports the values of the BD-BR metric for the test set used in this work.

Table 4. Bjontegaard-Delta Bit-Rate (BD-BR) with respect to High Efficiency Video Coding (HEVC).

Sequence	EVC	VVC
FlyingBirds	−21.96%	−23.90%
FortNite	−30.05%	−38.76%
CatRobot	−27.14%	−33.80%
DaylightRoad	−26.76%	−35.40%
SunsetBeach	−15.05%	−16.11%
ParkRunning	−23.24%	−12.19%
Average	**−22.29**%	**−26.69**%

Comparing the results in Table 3, based on the proposed method, and in Table 4, based on the BD-BR algorithm, for the BR differences of EVC and VVC with respect to HEVC, the offset between the two estimations is 0.24% for EVC and 1.63% for VVC. These results confirm that, even using a simple linear method, reliable values are obtained, with a small difference with respect to the BD-BR method.

The difference in computational complexity between the linear model described in this paper and the piecewise cubic model is not so significant when compared to the complexity of the video coding. In any case, the proposed method simplifies the computation, since it does not require the numerical integration step after the approximation step.

The two main advantages of the proposed method lie in the possibility to have direct information, both in numerical and in graphical terms, of the RD trend at different BR, both at the single sequence level and at the codec level. For a single sequence, the slope parameter of the RD model provides an immediate estimate of the RD variations from lower to higher bitrates, when comparing such a slope with another sequence encoded with the same codec, or the same sequence encoded with a different codec. Furthermore, with the model averaging, the averaged RD models of the different codecs allow a comparison of the performance of different codecs (in our case HEVC, EVC, VVC) when operating at lower or higher bitrates. Since the slopes of the averaged RD models for HEVC, EVC, VVC are not the same, the RD curves are not parallel to each other, and the numerical or graphical analysis gives an estimation of how the relative RD performance changes when changing the bitrate (or conversely, changing the PSNR for the inverse model (6)).

Vice versa, the typical BD-BR analysis just gives a numerical value for the differences in BR or PSNR between two different sets of encodings for a specific sequence. For the comparison between different codecs, the result is a single numerical value, obtained as an average over a number of sequences. So the different behavior of the codec at lower and higher Bitrates (or lower and higher PSNR) cannot be analyzed when comparing different sequences or when comparing with a different codec by the single numerical data resulting from the BD-BR method.

6. Conclusions

This paper presents a simple but reliable procedure to model and compare different video codecs under as similar test conditions as possible. Since it is well known that the PSNR vs BR characteristic is linear in a semi-logarithmic scale, this property is exploited to model the behavior of the codec for a single sequence and to average the obtained models over a number of sequences to have a general model for the codec. Finally, such averaged codec models can be compared to estimate the gain or loss in terms either of delta PSNR for a given BR range, or delta BR for a given PSNR range.

The main advantages of the proposed methodology in comparison to existing ones, and specifically the widely adopted BD-BR algorithm, are twofold. The BD-BR algorithm provides a single numerical value for the RD performance of a specific sequence encoded with a specific encoder. Such values are then averaged over several sequences to provide a single numerical value for the RD performance of a specific codec. With the proposed method, instead, we have a linear RD model for each sequence and an averaged RD model for a given codec. Such models allow the same comparison over the whole range of bitrates that can be done with the BD-BR algorithm. Furthermore, it is possible to study the behavior at lower or higher bitrates of the codecs, since the resulting RD curves can be more or less close to each other at different BR, depending on the slope coefficient. As an example, the RD curves of Figure 4 show graphically that the RD performance gain of VVC over EVC is larger at lower bitrates and smaller at higher bitrates. This comparison can be done also at the single sequence level, comparing the behaviour of the same codec over the different sequences. Besides graphically, all such comparisons can be performed also numerically, using the models as defined in (7) to (10).

The approach is presented for the rate distortion curves in terms of PSNR vs BR, but can easily be extended to other metrics such as VMAF and MOS.

References

1. Barnett, T.; Jain, S.; Andra, U.; Khurana, T. Cisco Visual Networking Index (VNI) Complete Forecast Update, 2017–2022. APJC Cisco Knowledge Network (CKN) Presentation. December 2018. Available online: https://www.cisco.com/c/dam/m/en_us/network-intelligence/service-provider/digital-transformation/knowledge-network-webinars/pdfs/1213-business-services-ckn.pdf (accessed on 24 September 2020).

2. Sullivan, G.J.; Ohm, J.R.; Han, W.J.; Wiegand, T. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *22*, 1649–1668. [CrossRef]

3. Mccann, K. MPEG-5 Essential Video Coding (EVC). In Proceedings of the ITU Workshop on "The Future of Media", Geneva, Switzerland, 8 October 2019.

4. Bross, B. Versatile Video Coding (VVC): The emerging new standard Versatile Video Coding. In Proceedings of the ITU Workshop on "The Future of Media", Geneva, Switzerland, 8 October 2019.

5. Ohm, J.R.; Sullivan, G.J.; Schwarz, H.; Tan, T.K.; Wiegand, T. Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC). *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *22*, 1669–1684. [CrossRef]

6. Dissanayake, M.B.; Abeyrathna, D.L.B. Performance Comparison of HEVC and H.264/AVC Standards in Broadcasting Environments. *J. Inf. Process. Syst.* **2015**, *11*, 483–494. [CrossRef]

7. Grois, D.; Marpe, D.; Mulayoff, A.; Itzhaky, B.; Hadar, O. Performance comparison of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC encoders. In Proceedings of the 2013 Picture Coding Symposium (PCS), San Jose, CA, USA, 8–11 December 2013; pp. 394–397. [CrossRef]

8. Barman, N.; Martini, M.G. H.264/MPEG-AVC, H.265/MPEG-HEVC and VP9 codec comparison for live gaming video streaming. In Proceedings of the 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), Erfurt, Germany, 31 May–2 June 2017; pp. 1–6. [CrossRef]

9. Grois, D.; Nguyen, T.; Marpe, D. Performance comparison of AV1, JEM, VP9, and HEVC encoders. In *Applications of Digital Image Processing XL*; Tescher, A.G., Ed.; International Society for Optics and Photonics (SPIE): Bellingham, WA, USA, 2018; Volume 10396, pp. 68–79. [CrossRef]

10. Katsavounidis, I.; Guo, L. Video codec comparison using the dynamic optimizer framework. In *Applications of Digital Image Processing XL*; Tescher, A.G., Ed.; International Society for Optics and Photonics (SPIE): Bellingham, WA, USA, 2018; Volume 10752, pp. 266–281. [CrossRef]

11. Bjøntegaard, G. Calculation of Average PSNR Differences between RD-Curves. In Proceedings of the Report VCEG-M33, ITU-T SG16/Q6, 13th Video Coding Experts Group (VCEG) Meeting, Austin, TX, USA, 2–4 April 2001.

12. Bjøntegaard, G. Improvements of the BD-PSNR Model. In Proceedings of the Report VCEG-AI11, ITU-T SG16/Q6, 35th Video Coding Experts Group (VCEG) Meeting, Berlin, Germany, 16–18 July 2008.

13. Chen, Y.; François, E.; Galpin, F.; Jullian, R.; Kerdranvat, M. Comparative Study of Video Coding Solutions VVC, AV1, EVC Versus HEVC. In Proceedings of the Report JVET-N0605, Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 15th Meeting, Gothenburg, Sweden, 3–12 July 2019.

14. Le Tanou, J.; Médéric, B. Analysis of Emerging Video Codecs: Coding Tools, Compression Efficiency and Complexity. In Proceedings of the SMPTE 2018 Annual Technical Conference and Exhibition at All-New Venue in Downtown Los Angeles, Los Angeles, CA, USA, 23–25 October 2018.

15. Schomaker, M.; Heumann, C. When and when not to use optimal model averaging. *Stat. Pap.* **2018**, *61*, 2221–2240. [CrossRef]

16. Hansen, B.E. Least Squares Model Averaging. *Econometrica* **2007**, *75*, 1175–1189. [CrossRef]