RATE-DISTORTION-COMPLEXITY ANALYSIS OF PARAMETRIC VIDEO CODECS

R. L. de Queiroz[†] *D. C. Garcia* [†] *Y.-H. Chen*[⋆] *R. Conceição*[+] *W.-H. Peng* [⋆] *L. V. Agostini*[+]

[†] Universidade de Brasilia, Brazil.
[⋆] Natl. Yang Ming Chiao Tung Univ., Taiwan.
[+] Universidade Federal de Pelotas, Brazil.

ABSTRACT

A parametric video coder-decoder (codec) is one where the performance is represented by a curve either in the plane or in higher-dimensional space. We propose a process for comparing parametric video compression systems through a rate-distortion-complexity (RDC) analysis. We discuss the popular 2D Bjontegaard delta metric and other correspondence-based curve-comparison methods, and their extensions to allow comparisons in RDC space. We argue that all these metrics in 3D may be innefective and we propose an absolute cost for each RDC curve. Such a metric involves the computation of Lagrangian costs of RDC points in the form $D+\lambda R+\gamma C$, and some integration. We also argue that any application may be associated to a (λ, γ) pair. Several state-of-the-art neural video codecs were compared, using the proposed metrics for the entire (λ, γ) plane for each codec through all the available dataset. We could, then, establish the best codecs for each application, i.e. for each (λ, γ) pair. Tests with 19 codecs (17 neural and 2 conventional video codecs) revealed that only 6 codecs would be of interest, each best suited for a range of applications.

1. INTRODUCTION

The evaluation of a video coder-decoder (codec) is typically made through rate-distortion (RD) analysis. There are, however, other factors involved with conceiving a cost function for a codec. Complexity (C) is perhaps the most important factor (cost) to be included apart from rate (R) and distortion (D). Other factors can be intelectual property (or lack thereof), manufacturing, operation and maintenance costs, or costs to rewrite the specs, replace the code base, replace equipments, etc. We are concerned here with a rate-distortion-complexity (RDC) analysis to evaluate the cost function of a codec.

Rate evaluation in video codecs is straightforward and so is distortion if it is measured in terms of mean squared error (MSE), The peak signal-to-noise ratio (PSNR) is a measure of quality derived from MSE. Other common distortion metrics are the Sum of Absolute Differences (SAD), primarily used in motion estimation [1, 2], and the Video Multi-Method Assessment Fusion (VMAF) [3]. Complexity, however, is not very well defined. Execution time is often used as a complexity metric and in standards or industry an idea that improves a codec can be rejected for being "too costly for so little gain", i.e. subjective decisions. We want such decisions to be taken using a quantitative RDC analysis. In order to illustrate the issue, consider the topmost plots in Fig. 1, which depict two typical RD curves for two codecs. It seems clear that one of the coders is best because it has lower rate and lower distortion than the other. However, if we rotate this plot in RDC space, in order to reveal the C axis, we get the bottom plot in Fig. 1. Then, the decision is not clear. We want to propose metrics to help in such a process.

Related to this work, Herglotz *et. al.* provided a comprehensive overview of the BD calculus and its existing implementations [4]. Katsenou *et. al.* [5] explore and compare energy consumption in both encoding and decoding processes across optimized, state-of-the-art traditional video codecs. Additionally, the authors propose the metric Energy-to-Bitrate (EBR). In another work [6], it was carried a comprehensive analysis of energy and compression efficiency across different video coding formats. An objective evaluation method named Generalized Bjøntegaard Delta PSNR (GBD-PSNR) was also proposed [7], which incorporates coding complexity as an additional dimension, resulting in a RDC evaluation surface. Such a metric, however, has some limitations. For example, when comparing two codecs with very different complexities. None of these works, however, consider neural codecs.

2. NOTES ON THE COSTS OF A CODEC

Complexity can be defined and characterized in many different ways. In traditional codecs, like H.264/AVC [8], H.265/HEVC [9], H.266/VVC [10], AV1 [11], etc., many works employ encoding and decoding runtimes as complexity metrics (see for example [12, 13, 14]) . However, using runtime to measure complexity has some circunstancial limitations, like platform dependence. For neural video codecs, an important metric is the number of multiply–accumulate (MAC) operations, usually expressed in thousands of MAC per pixel (kMAC/pel). This is an intrinsic complexity that is usually platform-independent. Other important metrics

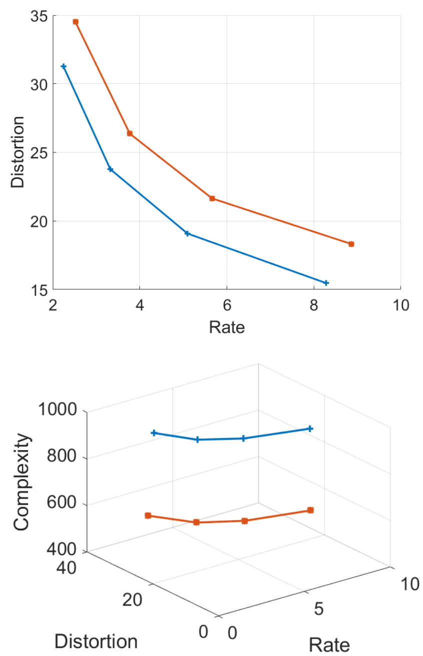

Fig. 1: Comparing two codecs by their performance curves. Top: in RD space. Bottom: in RDC space, which may introduce a different perspective.

for neural video codecs are the memory bandwidth, decoded picture buffer size, or power consumption. One could argue that complexity is a multidimensional problem. For the sake of the argument, we use here the intrinsic complexity in kMAC/pel, , instead of a subjective analysis of complexity, guided by computation time. We also adopt the usual linear combination of rate (R), distortion (D) and complexity (C). The analysis can be easily extended to other metrics.

We use the actual MSE distortion instead of PSNR (quality) in order to equally treat all three RDC dimensions. That is, we aim for minimizing the RDC cost along the three (or even more) non-negative dimensions.

A typical codec cost metric is the Bjontegaard delta (BD) [15]. This metric is commonly calculated either as rate savings (BD-rate), expressed in percentage terms, or as an improvement in PSNR (BD-PSNR), which is measured in decibels (dB). An important problem with BD is that it can lead to conflicting conclusions in many cases, as in the the example in Fig. 2, where BD-rate would favour one codec while BD-PSNR would favor the other.

The problem arises from the idea that comparing curves implies corresponding points in the curves. If one wants to compare two curves (C1 and C2) in RD space, assume the curves are parameterized as $(R(t), D(t))$, where t is a parameter for the curves from 0.0 to 1.0. Implicitely, when comparing points within a range of C1, every point t_1, in it, bears correspondence to a point t_2 in C2, from which we compute differences as $\delta R(t_1) = R_1(t_1) - R_2(t_2)$ and $\delta D(t_1) = D_1(t_1) - D_2(t_2)$. The integration over t_1 yields the error vector $(\Delta R, \Delta D)$. The BD-PSNR (actually BD-MSE) is the case where both points in C1 and C2 lie at the same R coordinate, whereas BD-rate is the same in regards to the D coordinate. A few other correspondences may be useful, such as *parametric* (where $t_1 = t_2$), *nearest* (pick the nearest point in C2 to the reference point in C1), and λ-oriented (corresponding points in C1 and C2 are aligned along a line with inclination determined by a parameter). These correspondences for a point in curve C2 are illustrated in Fig. 3.

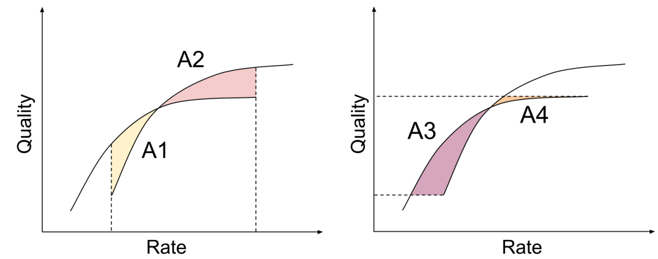

Fig. 2: Illustration of conflicting BD-PSNR and BD-rate for crossing RD curves.

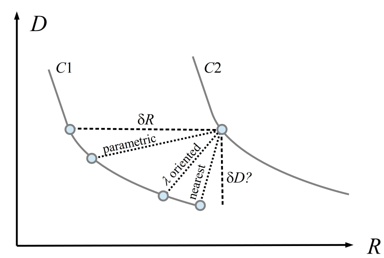

Fig. 3: Comparison of correspondence points in curve C1 to a point in C2.

The comparison yields $(\Delta R, \Delta D)$. Note that BD metrics set one of them to zero. If both ΔR and ΔD are negative (positive) it means that codec 1 is better (worse) than codec 2. Otherwise, the choice of which coder is better depends on the application. If we want to minimize $J = D + \lambda R$, it is expected that codec 1 would be better than codec 2 if $\Delta D < -\lambda \Delta R$. What is very important for this work is the fact that λ depends on the application. Furthermore, if $J = D + \lambda R$ is the only equation one minimizes, λ is the "application parameter".

Correspondence in between points may not make much sense in many cases. If the two curves have very different shapes, or are in very different positions, or have very different lengths, parametric or nearest-point correspondence may be less effective. In RDC space, with an extra dimension, these differences may be much more pronounced. If the finite curves, C1 and C2, are not close to each other, there may be no correspondences in ΔD or ΔR computation. If the curves are very different one may even find it difficult to determine which extremity corresponds to $t = 0$.

An alternative for comparing parametric curves C1 and C2 in RDC space may be to project the curves onto 2D planes (e.g. RD, RC, DC planes, or some application-oriented plane) and to use BD metrics with any of the above correspondence metrics. We do not discuss such method here, which we believe to be less effective than the one we are proposing.

3. ABSOLUTE COSTS IN RDC SPACE

The conclusion from the previous section is that, although there may be a few alternatives to compare two arbitrary RDC curves by comparing their geometry, none of the methods is robust. It may be more interesting to compute a cost for each curve and then compare curve costs.

In an RDC volume, assume non-negative RDC values, positive constants λ and γ, such that we want to minimize $J = D + \lambda R + \gamma C$. The (λ, γ) pair not only indicates the relative importance of distortion and complexity, but they are also a reflection of the application. Together, (λ, γ) span the "application space" and, for each aplication, one should find the (λ, γ) point that best suits the target application. We have indicated elsewhere how to calculate a suitable (λ, γ) point for a hypothetical streaming business application [28].

We can make a geometric illustration of the functions we are trying to minimize. In an RDC volume, assuming non-negative RDC values, and positive constants λ and γ, if one wants to minimize $J = D + \lambda R + \gamma C$ out of points $\{p_n\} = \{(r_n, d_n, c_n)\}$, the one that has the lowest $J_n = d_n + \lambda r_n + \gamma c_n$ is the one that is the closest to a *cost* plane

$$D + \lambda R + \gamma C = 0. \tag{1}$$

The distance z_i from point p_i to the cost plane is:

$$z_i = \frac{d_i + \lambda\, r_i + \gamma\, c_i}{\sqrt{1 + \lambda^2 + \gamma^2}}. \tag{2}$$

The coordinates (r'_i, d'_i, c'_i) of the point where p_i projects onto the plane are:

$$d'_i = d_i - q_i, \tag{3}$$
$$r'_i = r_i - q_i\lambda, \tag{4}$$
$$c'_i = c_i - q_i\gamma, \tag{5}$$

where

$$q_i = \frac{d_i + \lambda\, r_i + \gamma\, c_i}{1 + \lambda^2 + \gamma^2}. \tag{6}$$

A codec is defined here by a set of coding instantiations, or points in RDC space, forming a curve. The distance of each point to the plane is its cost. We want to measure the average distance of each curve to the plane, as illustrated in Fig. 4. For every pair of adjacent points p_i, p_j in this segmented curve of straight lines, the area forms a right trapezium. As such, the

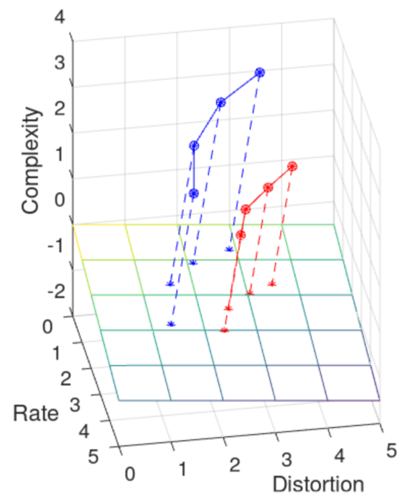

Fig. 4: Projection of the RDC points of two hypothetical codecs onto the plane $D + 2R + 10C = 0$.

average distance of a segment in the curve to the plane is the average of the distances from p_i, p_j to the plane. Since we have segments with different lengths one may weight each segment average by its respective projected length. In symbols, we let the average distance (to the plane) of points p_i and p_j be

$$z_{i,j} = \frac{z_i + z_j}{2}, \tag{7}$$

and define the length of the projected segment as

$$\ell_{i,j} = \sqrt{(d'_i - d'_j)^2 + (r'_i - r'_j)^2 + (c'_i - c'_j)^2}. \tag{8}$$

Then, the cost of the entire curve of N points is

$$J = \frac{\sum_{i=1}^{N-1} \ell_{i,i+1} z_{i,i+1}}{\sum_{i=1}^{N-1} \ell_{i,i+1}}. \tag{9}$$

Hence, in order to compare two RDC curves, for a given application point, we propose to compute the cost of each curve and, then, to compare the costs.

Of course, it depends on the application parameters λ and γ. We can, however, calculate (estimate) the cost J for all (λ,γ) points. For this, we can choose a number of (λ,γ) points, calculate their costs J and interpolate the rest. The result is a surface in the application plane $J(\lambda, \gamma)$ which indicates the best-suited applications for the given codec, i.e. the range of (λ,γ) values for which the given codec performs well.

We can, however, go further. Let there be N_c codecs (curves in RDC space) to test. For the m-th codec we can build $J_m(\lambda, \gamma)$. We, then, compare all the $J_m(\lambda, \gamma)$ for all the coders for all (λ,γ), i.e. we compute

$$B(\lambda, \gamma) = \arg\min_i J_i(\lambda, \gamma), \tag{10}$$

Table 1: Evaluated codecs.

i	codec	Decoding complexity (kMACs/pixel)
0	CANF-VC [16]	1748
1	DCVC [17]	762
2	DCVC-TCM [18]	924
3	DCVC-HEM [19]	1252
4	DCVC-DC [20]	924
5	DCVC-FM [21]	878
6	MaskCRT [22]	767
7	C16 [23]	541
8	C32 [23]	592
9	C64 [23]	762
10	CR16 [23]	541
11	CR32 [23]	593
12	CR64 [23]	764
13	MCR16 [23]	598
14	MCR32 [23]	649
15	MCR64 [23]	821
16	HyTIP [24]	873
17	H.265/HEVC [9]	≈ 1
18	H.266/VVC [10]	≈ 1.8

which indicates the best codec, among all candidates, for any application point. A profound result of examining B is that only the codecs whose indices are in the B map would have any sense pursuing (assuming the proposed metrics). Of course, different complexity metrics or different cost formulas would yield different results.

4. COMPARING CODECS

We have analyzed data comprised of four RDC points of 17 neural video codecs, including several state-of-the-art codecs and mainstream coding frameworks that are based on conditional coding, conditional residual coding, and masked conditional residual coding, as shown in Table 1, along with two standard video codecs (HEVC and VVC). Each of the 17 neural video codecs (except DCVC-HEM, DCVC-DC, and DCVC-FM) is trained to support four different rates with 4 distinct sets of network weights. Unlike the other neural video codecs, DCVC-HEM, DCVC-DC, and DCVC-FM are variable-rate codecs; they are able to compress an input video for a continuous range of bit rates. For these codecs, we produce 4 distinct bit rates corresponding to 4 quality levels. There are, thus, 76 coding instantiations to appreciate, generating 76 (r_{ij}, d_{ij}, c_{ij}) points, from which the 19 costs for the respective 19 codecs are computed according to Eq. (9).

The tests were carried to compress all the sequences in the test sets indicated in Table 2. There are 55 sequences, each with a different bit-depth, resolution, frame-rate, and number of frames. For each sequence and codec, rate is calculated in megabits per second (Mb/s). In order to average

Table 2: Test datasets.

Dataset	Number of sequences	Resolution
UVG [25]	6	1920×1080
HEVC-B [26]	5	1920×1080
HEVC-C [26]	4	832×480
HEVC-E [26]	3	1280×720
HEVC-RGB	6	1920×1080
MCL-JVC [27]	30	1920×1080

R across codecs, each rate value is normalized by converting it to bits/pixel and then denormalizing for a 1920×1080-pixel 30 Hz sequence. It is worth noting that rate R for each codec and configuration is computed as the average of all the 55 normalized rates, in Mb/s. Each of the 64 coding instantiations demand encoding all the testsets. Distortion D for each codec is computed as MSE (8-bit/color/pixel, i.e. 0-255 range) in the RGB domain, since not all neural video codecs support YUV coding. For 12-bit/color/pixel sequences, MSE is divided by 256 to make it compatible with the MSE of 8-bit sequences. Finally, D for a given codec and configuration is computed by averaging all the 55 distortions in MSE for each sequence. To quantify the complexity C of these neural codecs, we chose the decoding complexity in kMACs/pixel, which is intrinsic to their decoder architectures, as shown in Table 1. Note that similar evaluations can be made using the encoding complexity. We have included two non-neural codecs (HEVC and VVC) in the list. Their complexity is not intrinsic, but data-dependent. The decoder is not affected by motion estimation and mode decision, so that the worst-case complexity is related to a case using the largest size transforms for the entire image. It is hard to estimate thse numbers but members of the video codec industry estimated worst-case complexity to be in the neighbourhood of 1 KMAC/pel for HEVC. We then "guessed" worst-case complexities of 1 and 1.8 KMAC/pel for HEVC and VVC, respectively.

The RDC curves for all codecs tested are depicted in Fig. 5, from which not much visual information is ascertained without careful examination under multiple views. The figure also shows a typical view along the C axis, i.e. an RD plot, or a plot on the plane $\gamma = 0$, wherein complexity is unimportant. We need, thus, to compute our proposed metrics. For each codec, we computed $J_m(\lambda, \gamma)$ and spanned the values of λ and γ. An example of $J_m(\lambda, \gamma)$ for the CANF-VC coder is shown in Fig. 6, so the reader can have an idea of how it looks. We, then, repeated for all m, i.e. all codecs, and calculated $B(\lambda, \gamma)$, which is depicted in Fig. 7. In the $B(\lambda, \gamma)$ map in Fig. 7, we have also indicated the point $\lambda = 7.02$ and $\gamma = 1.14$ which can be shown to be reasonable parameters for a streaming application [28]. These interesting results point to the 6 codecs that would be the best we have tested, for our RDC criteria. Their choice depends on the application, i.e. on (λ, γ).

The traditional RD metrics corresponds to the bottom line

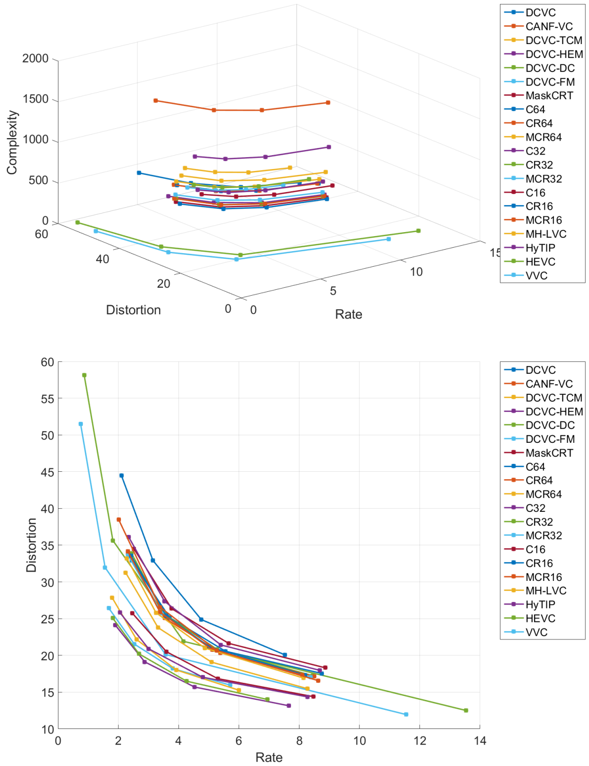

Fig. 5: RDC plot results for all 19 codecs tested. Bottom: RD plots, i.e. the view along the C axis, or the plane for $\gamma = 0$.

of the graphic, which would point to the HyTIP codec for lower rates, stepping through DCVC-DC and DCVC-FM as the rate increases. Very high rates are dominated by VVC perhaps because it has the highest rate points. As the importance of complexity increases, MaskCRT has a sweet spot of dominance, above which the whole application plane is dominated by non-neural coders. HEVC dominates at the top, higher γ, where lower complexity is more important. VVC dominates the rest. The results indicate that VVC would be preferred for the streaming application. Since the (λ, γ) numbers are in log domain, and many parameters were "guessed" we risk to say that the streaming application example may lie near the border in between HEVC and VVC. We expect the results for VVC and HEVC to be even better, if we carried the PSNR in YUV space rather than RGB.

5. CONCLUSIONS

We were motivated by the needs to also analyze the complexity of video codecs when evaluating their performances. This is particularly more relevant with the rapid development of complex neural video codecs. We recognize that the complexity analysis of a codec is a multifaceted problem, but we propose to use a simple complexity metric and a linear Lagrangian combination of rate (R), distortion (D) and complexity (C). We discussed why comparing curves would not be a good idea and proposed to construct absolute costs for each codec (a cost for the whole curve in RDC space). Costs were evaluated and compared for several modern neural video codecs, for all application points, allowing us to select the best codecs for all relative weights of RDC, in essence, for every application point. One has to find the (λ', γ') for the intended application, and, if the C metric is agreed upon, it is simple to find $B(\lambda', \gamma')$ and to chose the best video codec in an RDC sense.

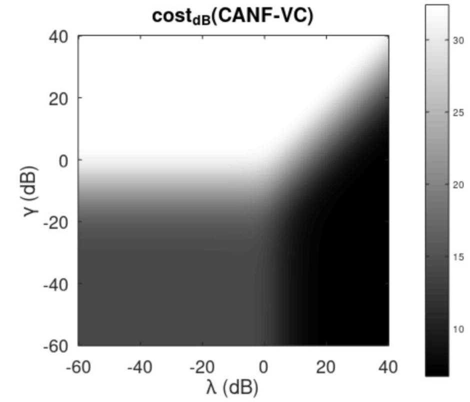

Fig. 6: The cost $J_m(\lambda, \gamma)$ for the CANF-VC codec. (λ, γ) are in log-domain (dB).

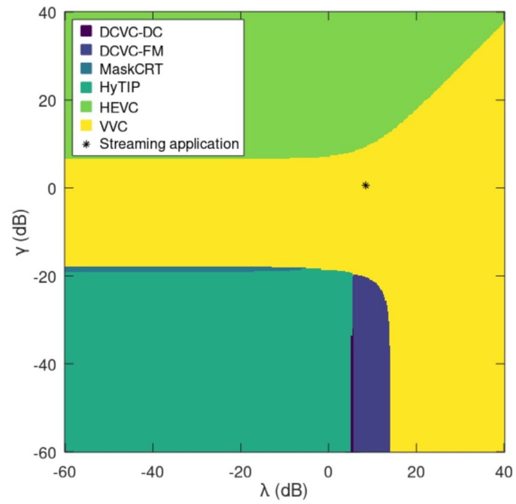

Fig. 7: $B(\lambda, \gamma)$, the map with the best codecs for all application points (λ, γ) for all 17 codecs tested. (λ, γ) are in log-domain (dB).

6. REFERENCES

[1] B. Bross, Y.-K. Wang, Y. Ye, S. Liu, J. Chen, G. J. Sullivan, and J.-R. Ohm, "Overview of the versatile video coding (VVC) standard and its applications," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 31, no. 10, pp. 3736–3764, 2021.

[2] J.-R. 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 Trans. Circuits and Systems for Video Tech.*, vol. 22, no. 12, pp. 1669–1684, 2012.

[3] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, and M. Manohara, "Toward a practical perceptual video quality metric," Netflix Tech Blog, 2016, [Online].

[4] C. Herglotz, H. Och, A. Meyer, G. Ramasubbu, L. Eichermüller, M. Kränzler, F. Brand, K. Fischer, D. T. Nguyen, A. Regensky *et al.*, "The Bjøntegaard Bible: Why your way of comparing video codecs may be wrong," *IEEE Trans. Image Proc.*, 2024.

[5] A. Katsenou, J. Mao, and I. Mavromatis, "Energy-rate-quality tradeoffs of state-of-the-art video codecs," in *Picture Coding Symp. (PCS)*, pp. 265–269, 2022.

[6] M. Kränzler, C. Herglotz, and A. Kaup, "A comprehensive review of software and hardware energy efficiency of video decoders," *arXiv:2402.09001*, 2024.

[7] X. Li, M. Wien, and J.-R. Ohm, "Rate-complexity-distortion evaluation for hybrid video coding," in *2010 IEEE Intl. Conf. Multimedia and Expo*pp. 685–690, 2010.

[8] ITU-T and ISO/IEC, "Advanced video coding for generic audiovisual services," ITU-T Recommendation H.264 and ISO/IEC 14496-10, 2003, version 2019.

[9] ——, "High efficiency video coding," ITU-T Recommendation H.265 and ISO/IEC 23008-2, 2013, version 2020.

[10] ——, "Versatile video coding," ITU-T Recommendation H.266 and ISO/IEC 23090-3, 2020.

[11] A. for Open Media, "Av1: Aomedia video codec," Available online: `https://aomedia.org/av1/`, 2018.

[12] Y.-H. Huang, T.-S. Ou, and H. H. Chen, "Fast decision of block size, prediction mode, and intra block for h.264 intra prediction," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 20, no. 8, pp. 1122–1132, 2010.

[13] M. Xu, T. Li, Z. Wang, X. Deng, R. Yang, and Z. Guan, "Reducing complexity of HEVC: A deep learning approach," *IEEE Trans. on Image Proc.*, vol. 27, no. 10, pp. 5044–5059, 2018.

[14] G. Correa, P. A. Assuncao, L. V. Agostini, and L. A. da Silva Cruz, "Fast HEVC encoding decisions using data mining," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 25, no. 4, pp. 660–673, 2015.

[15] G. Bjøntegaard, "Calculation of average PSNR differences between RD curves," in *ITU-T SG16/Q6, 13th VCEG Meeting, Austin, TX, USA, April 2001*, 2001.

[16] Y.-H. Ho, C.-P. Chang, P.-Y. Chen, A. Gnutti, and W.-H. Peng, "CANF-VC: Conditional augmented normalizing flows for video compression," in *Computer Vision – ECCV*, 2022.

[17] J. Li, B. Li, and Y. Lu, "Deep contextual video compression," in *Advances in Neural Information Processing Systems*, M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Eds., vol. 34. Curran Associates, Inc., 2021, pp. 18 114–18 125.

[18] X. Sheng, J. Li, B. Li, L. Li, D. Liu, and Y. Lu, "Temporal context mining for learned video compression," *IEEE Trans. on Multimedia*, 2022.

[19] J. Li, B. Li, and Y. Lu, "Hybrid spatial-temporal entropy modelling for neural video compression," in *Proc. of the 30th ACM Intl. Conf. on Multimedia*, 2022.

[20] ——, "Neural video compression with diverse contexts," in *Proc. IEEE/CVF Conf. on Comp. Vision and Pattern Recognition*, CVPR, Vancouver, CA, June 2023.

[21] ——, "Neural video compression with feature modulation," in *Proc. IEEE/CVF Conf. on Comp. Vision and Pattern Recognition*, CVPR, Seattle, WA, USA, June 2024.

[22] Y.-H. Chen, H.-S. Xie, C.-W. Chen, Z.-L. Gao, M. Benjak, W.-H. Peng, and J. Ostermann, "MaskCRT: Masked conditional residual transformer for learned video compression," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 34, no. 11, pp. 11 980–11 992, 2024.

[23] Y.-H. Chen, K.-W. Ho, M. Benjak, J. Ostermann, and W.-H. Peng, "On the rate-distortion-complexity trade-offs of neural video coding," in *Proc. 26th IEEE Intl. Workshop on Multimedia Signal Proc. (MMSP)*, 2024, pp. 1–6.

[24] Y.-H. Chen, Y.-C. Yao, K.-W. Ho, C.-H. Wu, H.-T. Phung, M. Benjak, J. Ostermann, and W.-H. Peng, "Hytip: Hybrid temporal information propagation for masked conditional residual video coding," *arXiv:2508.02072*, 2025.

[25] A. Mercat, M. Viitanen, and J. Vanne, "UVG dataset: 50/120fps 4K sequences for video codec analysis and development," in *Proc. ACM Multimedia Syst. Conf.*, Istanbul, Turkey, June 2020.

[26] F. Bossen *et al.*, "Common test conditions and software reference configurations," *JCTVC-L1100*, vol. 12, no. 7, 2013.

[27] H. Wang *et al.*, "MCL-JCV: A JND-based H.264/AVC video quality assessment dataset," in *Proc. IEEE Intl. Conf. Image Proc.*, ICIP, Phoenix, AZ, USA, 2016, pp. 1509–1513.

[28] R. L. de Queiroz, D. Garcia, Y.-H. Chen, R. Conceicao, W.-H. Peng, and L. V. Agostini, "Application space and the rate-distortion-complexity analysis of neural video codecs," arXiv:2509.05929, 2025.