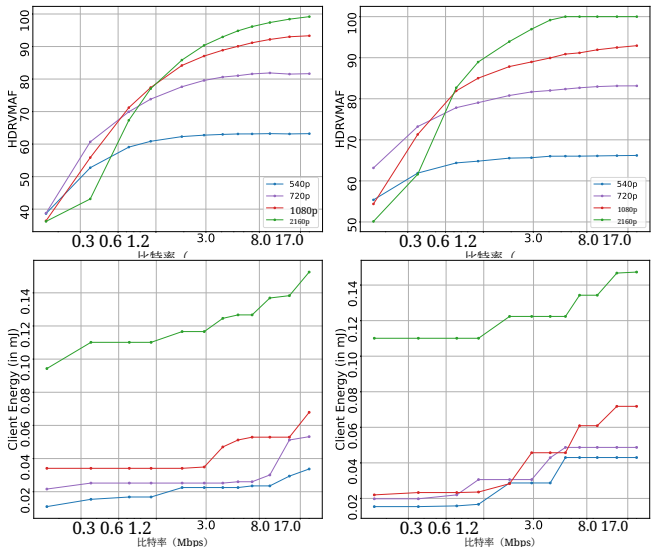
面向高动态范围自适应视频流的码率-质量-能耗凸包

Amritha Premkumar*、Christian Herglotz*
*科特布斯-森夫滕贝格勃兰登堡工业大学计算机工程系，德国

arXiv:2611.28817v1 [eess.IV] 21 Nov 2026

摘要—与标准动态范围（SDR）视频相比，高动态范围（HDR）视频的解码复杂度和能耗显著更高，但大多数自适应流媒体流程仍然依赖采用面向 SDR 的指标（如 VMAF）优化的码率阶梯。这些阶梯无法捕捉 HDR 特有的感知特性，也忽略了解码器侧能耗，因而产生能效低下的工作点。本文提出一种码率-质量-能耗凸包优化框架，综合考虑比特率、HDR 感知质量（HDRVMAF）以及客户端处理能耗，包括解码和升频。我们在 LIVE HDR 视频质量评估数据库中的 12 段 4K HDR 序列上开展实验，每段视频评估 72 种表示。结果表明，传统的 VMAF 优化码率阶梯经常会选择高分辨率编码，而在计入能耗后，这些编码往往不再占优。与默认 HLS 码率阶梯相比，我们的 HDR 感知方法在 HDRVMAF 相同的情况下，最高可节省54.34% 的比特率；在更高的能耗加权下，客户端能耗相对最多降低43.10% 。这些结果证明，所提出的框架可作为构建 HDR 码率阶梯的实用基础，明确呈现并控制感知质量-能耗权衡。

索引词—**高动态范围、自适应视频流、率–质量–能耗优化、能耗感知视频解码。**



(a) *CourtYard* (b) *NightStreet*

图1：LIVE HDR视频质量评估数据库中具有代表性的HDR序列在不同分辨率下的码率-质量和码率-能耗曲线。

I. 引言

与标准动态范围（SDR）内容相比，高动态范围（HDR）视频具有更宽的亮度范围、更高的峰值亮度和更广的色域，因此已成为现代流媒体服务的重要组成部分。这些特性提升了视觉真实感，但也增加了编码和解码的复杂度，导致客户端设备的能耗更高，尤其是在移动平台和采用功耗受限的片上系统（SoC）架构的智能电视上 [1]。因此，HDR 流媒体系统不仅必须优化感知质量，还必须提高播放能效。图 1 展示了比特率与 HDR 感知质量以及典型 HDR内容的客户端处理能耗。更高的分辨率可获得更优的 HDRV-MAF [2]分数，但需要消耗显著更多的客户端处理能耗；较低的分辨率则能降低能耗，但会牺牲感知质量。

自适应流媒体通过以多种分辨率和比特率对内容进行编码来应对带宽变化，使客户端能够选择合适的表示。传统的比特率梯级构建依赖于使用 VMAF 等面向 SDR 的指标生成的率失真凸包 [3]。尽管这些方法能够有效提升压缩效率，但它们假设比特率和感知质量是决定最优工作点的唯一因素。实际上，不同表示在播放期间会产生不同的计算成本。高分辨率和高位深视频由于像素数量更多、内存带宽需求更高以及预测结构更复杂，因此需要更高强度的解码 [4], [5]。相反，低分辨率流需要进行空间升频，这同样会产生不可忽略的能耗。对于 HDR视频而言，这些影响尤为明显，因为更高的位深和 HDR专用编解码工具会进一步增加解码器工作负载。因此，比特率梯级构建应同时考虑感知质量和客户端处理能耗，而后者会因编码表示的不同而变化。

关于逐标题和内容自适应编码的现有研究表明，通过多分辨率表示设计和凸包剪枝，可以显著提升压缩效率。Netflix 等业界采用的逐标题优化方法 [6]以及采用峰值信噪比、结构相似性和 VMAF 的学术方法 [7]–[9]均表明，内容依赖型梯级构建能够提高率失真效率。然而，这些方法通常假设不同表示的解码器成本一致，并忽略了

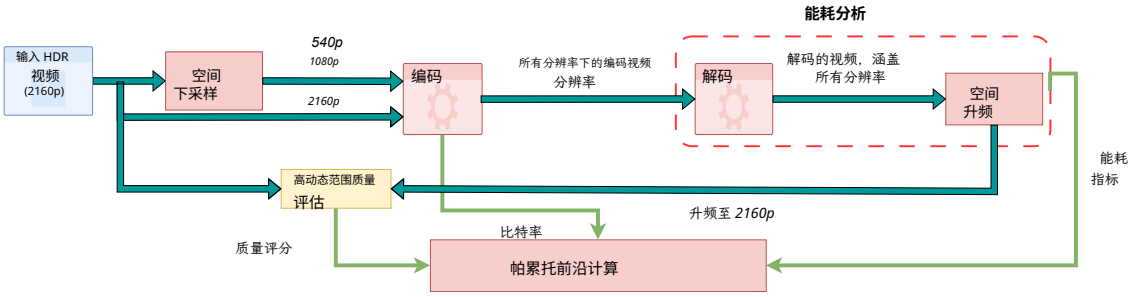


图 2：所提出的 HDR 流媒体码率-质量-能耗优化框架概览。

分辨率自适应带来的能耗影响。能耗感知流媒体研究考察了编码器端能耗优化 [1], [10] ，或将解码器复杂度作为功耗的替代指标 [11],[12]。尽管这些研究提供了有价值的见解，但它们要么优化编码器端处理，要么将解码器复杂度用作播放成本的间接指标。然而，在比特率梯级构建过程中，它们很少联合考虑实测的客户端处理能耗和感知质量，也未将空间升频带来的能耗影响纳入考量。

另一项重要挑战是 HDR 专用感知质量评估。传统指标（如峰值信噪比和结构相似性）无法充分建模 HDR 亮度适应、亮度掩蔽或感知阈值 [13]。多种面向 HDR 的指标，包括 HDR 视觉差异预测器 (HDR-VDP)-3 [14], HDR 视频质量度量 (HDR-VQM) [15], 统一质量评估器融合 (FUNQUE+) [16], HDRPatchMAX [17], 和显示无关梯度幅值相似性 (DIGMS) [18],, 都能更好地契合主观感知。其中，HDRVMAF [2] 已成为一种性能出色的全参考指标，并被本研究用作主要质量度量。

综合来看，这些观察结果揭示了当前 HDR 流媒体研究中的一个空白：比特率梯级构建仍主要由码率-质量考量驱动，而客户端处理能耗通常被单独处理。因此，在表示选择过程中，HDR感知质量、解码能耗与升频成本之间的相互作用并未得到明确考虑。为弥补这一空白，我们提出了一种码率-质量-能耗凸包框架，用于 HDR 自适应流媒体。该方法将考虑 HDR 特性的感知质量和客户端处理能耗纳入统一的表示选择流程。

本工作的主要贡献如下：

- 我们提出了一种速率-质量-能耗优化框架，在比特率梯级构建过程中综合考虑比特率、HDR感知质量、解码器能耗和升频能耗。
- 我们制定了一种可调的基于效用的表示选择策略，用于揭示并控制感知质量与客户端处理能耗之间的权衡。
- 我们证明，一旦将客户端能耗纳入考量，传统质量驱动优化所选出的许多表示就会变得次优，并量化由此带来的比特率和能耗节省。

II. HDR 码率-质量-能耗建模

本节介绍所提出的用于表示选择的 HDR 码率-质量-能耗建模框架。每种候选表示均根据三个属性进行评估：（i）传输码率；（ii）HDR感知质量；（iii）客户端处理能耗，包括解码和可选的升频。这些属性共同定义了图 2 所示的优化空间。与仅优化比特率和质量的传统率失真模型不同，所提出的模型明确纳入播放能耗，从而能够同时适应视觉要求和设备级约束。

A. 表示评估

一种表示由其空间分辨率 $r \in \mathcal{R}$和目标比特率 $b \in \mathcal{B}$ 定义。每种表示都会在客户端解码；当 r 低于显示分辨率时，还会在质量评估前进行升频。升频操作采用通用方式处理，可根据目标播放系统，使用传统插值滤波器或学习式视频超分辨率方法实现。对于每一组（r, b）组合，框架都会记录：

- 比特率 $R(r, b)$：实际达到的编码比特率。
- HDR感知质量 $Q(r, b)$：通过 HDRV-MAF 比较解码后（以及必要时升频后）的输出与 HDR 参考视频来测量。
- 客户端处理能耗 $E_{\text{总计}}(r, b)$：解码和空间升频所需的绝对能耗。

联合评估这些量，可以捕捉压缩效率、感知质量与播放能耗之间的权衡，而传统的码率—质量分析往往忽略了这些权衡。

B. 能耗分解

客户端处理总能耗建模如下：

$$E_{\mathrm{tot}}(r,b) = E_{\mathrm{dec}}(r,b) + E_{\mathrm{up}}(r,b), \tag{1}$$

其中， E_{dec} 表示编解码器解码操作，而 E_{up}表示当编码分辨率低于显示分辨率时与空间升频相关的能耗。

这种分解将编解码器复杂度与后处理开销分离开来，使该框架能够确定播放能耗主要由解码还是缩放所主导。

C. 能耗-质量权衡的标量效用

为平衡感知质量与能效，定义标量效用函数如下：

$$J(r,b;\alpha_D) = Q(r,b) - \alpha_D E_{\mathrm{tot}}(r,b), \tag{2}$$

其中， $\alpha_D \geq 0$ 控制质量–能耗权衡。较小的取值优先考虑视觉质量，而较大的取值则会越来越多地惩罚能耗。 α_D 并非表示物理常数，而是一个权衡参数（等价地说，在受能耗预算约束的质量最大化问题的约束形式中，它是一个拉格朗日乘子）。

为提高不同内容和硬件平台之间的可比性，也可以将效用表示为归一化形式：

$$J'(r,b;\alpha_D) = \frac{Q(r,b) - Q_{\min}}{Q_{\max} - Q_{\min}} - \alpha_D \frac{E_{\mathrm{tot}}(r,b) - E_{\min}}{E_{\max} - E_{\min}}, \tag{3}$$

其中，($Q_{\min}$， $Q_{\max}$) 和 ($E_{\min}$， $E_{\max}$) 是在候选表示集合上计算得出的。

D. 通过效用最大化进行分辨率选择

对于每个目标比特率，通过最大化效用来选择最优空间分辨率：

$$r^*(b;\alpha_D) = \arg\max_{r\in\mathcal{R}} J(r,b;\alpha_D), \tag{4}$$

或者等价地，最大化式（3）中的归一化效用。由此得到的比特率到分辨率的映射构成了能耗感知比特率阶梯，综合考虑了 HDR 感知质量和客户端处理能耗。

由于效用取决于内容特征和 α_D，所选分辨率会自然适应不同的运行条件。优先考虑质量时，高细节内容倾向于采用更高分辨率；而随着 α_D 增加，选择会转向能耗更低的表示，从而实现设备感知和内容自适应的比特率阶梯。

III. 实验设置

数据集：评估采用了来自 LIVE HDR视频质量评估数据库 [], 的 12 段 4K 高动态范围视频序列，这些序列经过筛选，能够体现现代 HDR 流媒体内容中常见的多样空间纹理、运动特征和亮度分布。数据集涵盖低运动和高运动场景，以及明亮和暗淡的序列，确保内容具有广泛的多样性。所有视频均以原生 2160p 分辨率处理，同时保留整个编码、解码和质量评估流程中的 HDR 格式及元数据。

表 I：实验参数及取值。

参数	取值
$\mathcal{R}$	{ 360, 540, 720, 1080, 1440, 2160 }
$\mathcal{B}$	0.160 | 0.360 | 0.600 | 0.900 | 1.600 | 2.400
	3.400 | 4.500 | 5.800 | 8.100 | 11.600 | 16.800
α_D	[0, 1]
目标编码器	x265 [较慢][帧内周期=1秒]
目标解码器	FFmpeg 原生 HEVC 解码器

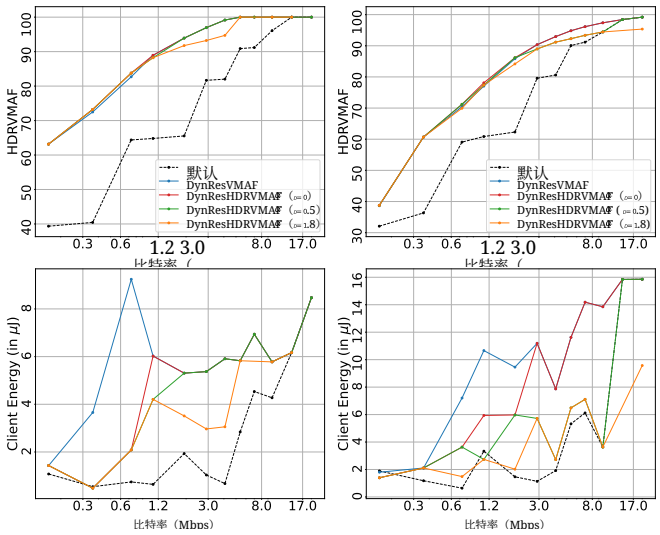
编码配置： 每个源序列均采用表 I 中列出的多种空间分辨率和比特率进行编码。所有编码均使用带有*slower* 预设的 x265 v4.0 编码器，并采用十位内部表示和输出。保持编解码器配置、GOP 结构和码率控制设置不变，可最大限度地减少混杂因素，因此观测到的感知质量和客户端处理能耗差异主要由分辨率和比特率引起。

感知质量评估： 感知质量使用峰值信噪比、VMAF [3],和 HDRVMAF [2], 进行评估，其中 HDRVMAF 由于对 HDR 特有失真具有更高的敏感性而作为主要指标。每个表示均使用 FFmpeg 进行解码，并在需要时使用 FFmpeg 中的双三次插值将其放大到 4K，然后与原始 HDR 参考视频进行比较。尽管商用设备通常使用硬件加速解码和缩放，但基于软件的流程为不同表示之间的相对比较提供了一致的环境。与面向 SDR 的指标相比，HDRVMAF 能更好地捕捉 HDR 的亮度、对比度和细节，因此能够提供更具感知意义的评估。

能耗测量： 客户端处理能耗在配备 8GB 内存的 Intel Core i5-1135G7 系统上进行测量。解码（以及升频）能耗测量五次，报告值取各次运行结果的平均值。能耗使用 CodeCarbon [19], 进行测量，该工具通过访问 Intel RAPL 计数器来估算中央处理器和动态随机存取存储器能耗。系统尽量减少后台服务，同时不计入显示和网络能耗。对于公式 (3) 中的归一化效用，使用固定的质量上下界 $Q_{\min} = 0$ 和 $Q_{\max} = 100$ ；由于优化依赖相对支配关系而非绝对能耗单位，因此直接使用测得的能耗值。

评估指标：使用 Bjøntegaard-Delta 指标 [20]评估比特率梯度的性能。Bjøntegaard Delta 码率值（BDR_{P}、BDR_{V}和 BDR_{H}）分别衡量在峰值信噪比、VMAF 和 HDRVMAF 下达到相同质量所需的平均比特率差异百分比。此外，Bjøntegaard Delta 客户端能耗（BDCE）量化了与基准比特率梯度相比，在维持相同 HDRVMAF 时客户端处理能耗的平均相对变化。负值表示相较于参考方法有所改进。

基准方法：将提出的框架与两种广泛使用的码率阶梯构建策略进行比较：



(a) *NightStreet* (E_Y:54.99, h:10.23, L_Y:81.49) (b) *CourtYard* (E_Y:103.08, h:12.96, L_Y:27.07)

图3：代表性HDR序列的码率-质量和码率-能耗曲线。

1) 默认：标准的苹果 HTTP实时流媒体 (HLS) 比特率阶梯，使用预定义的分辨率–比特率对，并作为基线表示集 [21]。 2) *DynResVMAF*：一种内容自适应方法，针对每种比特率选择解码并升频后 VMAF 最高的分辨率，从而在不受固定阶梯约束的情况下提供以质量为导向的上界 [6],[8]。

IV. 结果与分析

1) 码率-质量-能耗帕累托前沿： 图3展示了代表性HDR序列在二维码率-质量和码率-能耗空间中的表现。尽管最大化感知质量的配置通常会带来更高的客户端解码能耗，但其他工作点能够以显著更低的能耗实现相当的HDRVMAF。因此，在将解码能耗纳入优化目标后，一些在传统码率–质量空间中看似具有吸引力的配置将变为劣势配置。

2) 能耗加权的影响： 表II报告了不同能耗加权因子 α_D下所评估码率阶梯的平均性能。在 $\alpha_D = 0$下，优化器优先考虑HDRVMAF，因此倾向于选择高分辨率编码。增大 α_D 会使选定的表示转向低分辨率、更加节能的工作点，通常为1080p或720p。如表II所示，这一转变在保持显著比特率节省的同时降低了客户端处理能耗，证实能耗感知优化会显著影响码率阶梯设计。

3）客户端能耗与感知质量：图1和图3表明，随着比特率增加，客户端处理能耗急剧上升，而HDRVMAF的提升却逐渐趋于平缓。在多个序列中，较低分辨率达到的感知质量几乎与较高分辨率相同，但能耗显著更低，解码和升频成本也更低，因此在能耗约束下是很有吸引力的工作点。

表II：与默认相比，各比特率阶梯的平均结果。箭头表示性能变化趋势。

方案	α_D	BDR_P [%] ↓	BDR_V [%] ↓	BDR_H [%] ↓ ↑	BDCE [%] ↓
DynResVMAF	0	-20.23	-52.00	-42.04	123.04
DynResHDRVMAF	0	-20.15	-50.12	-54.34	27.15
	0.02	-16.80	-42.32	-47.60	-10.30
	0.04	-15.95	-41.84	-47.41	-11.86
	0.16	-11.20	-36.51	-42.21	-33.40
	0.32	-6.95	-28.62	-29.60	-43.10

4）内容依赖行为：内容特征会显著影响最优工作点（参见图3）。这些代表性序列分布在内容特征空间的不同区域：*NightStreet* 的空间纹理（E_Y: 54.99）和运动（h: 10.23 ）较少，但亮度较高（ L_Y: 81.49 ）；而 *CourtYard* 的空间纹理（E_Y: 103.08）和运动（h: 12.96）更丰富，亮度则显著较低（L_Y: 27.07） [22]。因此，细节更丰富、动态变化更明显的*CourtYard* 受益于更高分辨率的编码，从而保留感知质量；而画面更平滑的 *NightStreet*呈现出更平缓的HDRVMAF曲线，使较低分辨率的编码也能以更低的解码能耗实现相近的质量。这种差异表明，单一的内容无关码率阶梯并不足够，也促使我们采用针对不同内容的能量感知自适应策略。

V. 结论

本文提出了一种用于HDR视频流媒体的码率–质量–能耗凸包框架，联合建模比特率、HDR感知质量以及客户端解码和升频产生的处理能耗。通过在这一三维空间中评估候选表示，该框架揭示了传统码率–质量优化所忽略的权衡。在十二个4K HDR视频序列上的实验表明，许多在码率–质量域中看似最优的高分辨率编码，在考虑能耗后会变为劣势选择；HDR感知选择在保持等效感知质量的同时，可实现显著的比特率降低，其中 BDR_H 最低可达−54.34 %， BDR_V 最低可达 −50.12%。能耗感知优化还显著降低了相对客户端能耗，随着能耗权重增大，BDCE从 $\alpha_D = 0$处的27.15%提升至 −43.10%。因此，增加 α_D 会使最优工作点转向较低分辨率的表示，这反映出高分辨率下HDRVMAF趋于饱和，而处理能耗却快速增加。因此，所提出的框架能够通过由 α_D控制的可调质量–能耗权衡，实现设备感知HDR流媒体。

未来的工作包括将 α_D 适配到实时设备条件下，在硬件加速移动平台上验证该框架，并将其集成到基于 DASH 的流传输系统中，以实现实际的节能HDR传输。

参考文献

[1] C. Herglotz、S. L. Moan 和 A. Mercat，“HDR 视频编码中的能耗降低机会”，载于*2024* 年 *IEEE* 国际图像处理会议（*ICIP*），2024 年 10 月，第 3654–3660 页。[2] Z. Shang 等，“HDR 视频的主观与客观质量评估研究”，*IEEE* 图像处理汇刊，第 33 卷，第 42–57 页，2024 年。[3] Z. Li 等，“VMAF：旅程仍在继续”，载于*Netflix* 技术博客，2018 年 10 月。 [在线]。可访问：https://netflixtechblog.com/ vmaf-the-journey-continues-44 51 9 12 b ee ed G. J. Sullivan 等，“高[4]效视频编码（HEVC）标准概述”，*IEEE* 视频技术电路与系统汇刊，第 22 卷，第 12 期，第 1649–1668 页，2012 年。[5] A. Mercat 等，“我们需要 10 位吗？面向节能HDR视频流传输的 HEVC 编码器评估”，*IEEE* 电路与系统新兴及重点专题期刊，第 15 卷，第 1 期，第 31–43 页，2025 年 3 月。[6] J. De Cock 等，“云端基于复杂度的一致质量编码”，载于*2016* 年 *IEEE* 国际图像处理会议（*ICIP*），2016 年 9 月，第 1484–1488 页。[7] A. V. Katsenou、J. Sole 和 D. R. Bull，“与内容无关的自适应视频流码率阶梯预测”，载于*2019* 年图像编码研讨会（*PCS*）。IEEE，2019 年 11 月，第 1–5 页。[8] V. V. Menon 等，“用于自适应直播流传输的感知 JND 的两遍逐标题编码方案”，*IEEE* 视频技术电路与系统汇刊，第 34 卷，第 2 期，第 1281–1294 页，2024 年 2 月。[9] Vibhoothi 等，“在采用 AV1 的逐片段拉格朗日乘子优化中对 HDR 质量指标的比较”，载于*2023* 年 *IEEE* 多媒体与博览会国际会议（*ICME*），2023 年 7 月，第 1655–1660 页。[10] V. V. Menon 等，“结合感知延迟的动态分辨率编码实现节能自适应视频流”，载于第三届 *Mile-High* 视频会议论文集，2024 年，第 21–27 页。[11] Z. Azimi 等，“面向自适应 VVC 流传输的感知解码复杂度的码率阶梯估计”，载于*2024* 年第 *32* 届欧洲信号处理会议（*EUSIPCO*），2024 年，第 942–946 页。

[12] P. T. Rajendran等，“面向自适应视频流的能耗与质量感知可变帧率帕累托前沿”，载于*2024*年*IEEE*国际视觉通信与图像处理会议（*VCIP*），2024 年 12 月，第 1–5 页。[13] Z. Shang等，“不同环境条件下高动态范围视频的主观评估”，载于*2022*年*IEEE*国际图像处理会议（*ICIP*），2022年10月，第 786–790 页。[14] R. K. Mantiuk、D. Hammou 和 P. Hanji，“HDR-VDP-3：一种用于预测高动态范围和常规内容中的图像差异、质量及对比度失真的多指标”，2023年。 [在线]。网址：https://arxiv.org/abs/ 2304 13625 。M. Narwaria等，“HDR-VQM：一种用于高[15]动态范围视频的客观质量度量”，《信号处理：图像通信》，第 35 卷，第 46–60 页，2015年。[16] A. K. Venkataramanan等，“一种用于压缩高动态范围视频质量评估的 FUNQUE 方法”，载于*2024*年图像编码研讨会（*PCS*），2024年6月，第 1–5 页。[17] J. P. Ebenezer等，“HDR 还是 SDR？缩放和压缩视频的主客观研究”，《IEEE 图像处理汇刊》，第 33 卷，第 3606–3619 页，2024 年。[18] K. Zhang等，“一种面向高动态范围图像的显示无关质量评估方法”，《IEEE信号处理快报》，第 29 卷，第 464–468 页，2022年。[19] B. Courty等，“mlco2/codecarbon：v2.4.1”，2024年5月。 [在线]。网址：https://doi.org/ 10 5281 11171501 。/zenodo。G. Bjontegaard，“[20]RD 曲线之间平均峰值信噪比差异的计算”，*VCEG-M33*，2001年。[21] 苹果公司，“Apple 设备的 HLS 编写规范”。 [在线]。网址：https://developer.apple.com/documentation/http-live-streaming/hls-authoring-specification-for-apple-devices

[22] V. V. Menon等，“自适应视频流高效编码的绿色视频复杂度分析”，载于首届绿色多媒体系统国际研讨会论文集，2023年，第 16–18 页。[在线]。网址：https://doi.org/10.1145/3593908.3593942