通用视频编码中用于码率控制的内容驱动帧级比特预测

Amritha Premkumar*、Prajit T Rajendran†、Vignesh V Menon*‡、Christian Herglotz*
*勃兰登堡科特布斯-森夫滕贝格工业大学计算机工程教席，德国
†法国巴黎萨克雷大学、法国原子能和替代能源委员会、List，F-91120，帕莱索，法国
‡弗劳恩霍夫海因里希·赫兹研究所视频通信与应用部，德国柏林

arXiv:2602.06242v1 [eess.IV] 5 Feb 2026

摘要——码率控制通过在各帧之间高效分配比特，在满足目标比特率的同时保持质量。通用视频编码（VVC）中的传统双遍码率控制（双遍码率控制）依赖解析码率–QP模型，而这类模型往往无法捕捉非线性时空变化，导致质量不稳定，并因多次试编码而增加复杂度。本文提出一种内容自适应框架，在随机森林回归中利用视频复杂度分析器（VCA）的轻量级特征和量化参数来预测帧级比特消耗。在采用 VVenC 编码的超高清序列上，该模型与真实值高度相关，I 帧、P 帧和 B 帧的 R^2 值分别为 0.93、0.88 和 0.77。将其集成到码率控制环路后，该模型实现了与双遍码率控制相当的编码效率，同时将总编码时间缩短了 33.3%。结果表明，由 VCA 驱动的比特预测是一种计算高效且准确的替代方案，可取代传统的码率–QP模型。

索引词—**视频编码、码率预测、随机森林、视频复杂度分析、码率控制、自适应流媒体。**

I. 引言

随着流媒体平台、移动设备和互联网电视（OTT）服务的扩展，对高质量视频内容的需求持续增长 [1]。超高清（UHD）格式、高帧率以及虚拟现实等沉浸式应用，进一步增加了视频传输系统的复杂性。在此背景下，高级视频编码（AVC） [2], 高效视频编码（HEVC） [3], 和通用视频编码（VVC） [4]等视频编码标准在实现高效压缩和传输方面发挥着核心作用。视频编码的一项关键挑战是码率控制，即在满足目标比特率限制的同时，在各帧和编码单元之间分配比特，以保持感知质量 [5]。准确的比特预测是实现稳定码率控制以及高效存储和传输的基础。

传统的码率控制算法通常依赖简化的解析模型，将量化参数（QP）与输出码率关联起来 [5]。这些模型通常假设 QP 与比特数之间存在线性或指数关系；然而，实际的比特消耗很大程度上取决于内容特征。经典的码率控制率失真（R–D）建模研究了参数分布和帧层模型 [6]–[8], ，以及针对实时和低延迟运行的方案 [9]–[12]。由于比特消耗会随空间细节和运动而大幅变化，启发式模型往往难以泛化，因此常需采用保守的缓冲区控制或试编码来避免超调，从而增加计算成本和能耗。近期的双遍编码和快速首遍编码设计通过下采样和一致性质量控制缓解了这些问题 [13]–[15]。

为解决这一局限，研究日益聚焦于内容自适应码率控制。国际电信联盟（ITU）推荐的空间信息（SI）和时间信息（TI）等复杂度度量，已被用于预测编码行为 [16]。然而，它们与码率或编码时间等编码结果的相关性有限，尤其是在现代编解码器中 [17]。近年来，机器学习方法已被用于建模内容特征与编码参数之间的关系，从而支持面向码率梯级构建、逐标题编码和码率控制的内容感知双遍优化 [18], [19]。

与近期依赖编码器内部特征或深度网络的学习型码率控制方法，以及专注于优化片段级编码分辨率、码率阶梯或能效的方法不同， [20], [21],本文如图 1 所示，采用源自开源视频复杂度分析器（VCA）的轻量级内容特征 [17] 来预测帧级比特消耗。VCA 使用基于块级离散余弦变换分析得出的频域描述符，对空间纹理能量、亮度和帧间变化进行量化。这些特征的计算开销较低，且与变换尺寸、预测模式选择等编码器决策高度相关，因此适用于实时码率估计。针对 I 帧、P 帧和 B 帧的不同编码行为，模型采用帧类型专属模型；这些帧对空间复杂度和时间复杂度的依赖各不相同。所提出的模型在第一遍中直接根据 VCA 特征和量化参数估算帧级比特消耗，从而无需额外的分析遍次即可实现合理的码率分配和缓冲区控制。这缩小了单遍方案与双遍方案之间的差距，在保持单遍延迟的同时达到接近双遍码率控制（2pRC）的准确度，并为将该方法实际集成到 VVenC 的码率控制环路中提供了可行途径。

我们在使用 VVenC [22] 编码的大型超高清数据集上进行评估，结果表明，所提出的模型具有很高的准确性，I 帧、P 帧和 B帧的决定系数（R^2）分别为 0.93、0.88 和 0.77；同时，各帧类型的平均绝对百分比误差（MAPE）均保持较低水平，这证实了 VCA 特征能够有效捕捉影响码率消耗的视频复杂度。

本文的贡献如下：

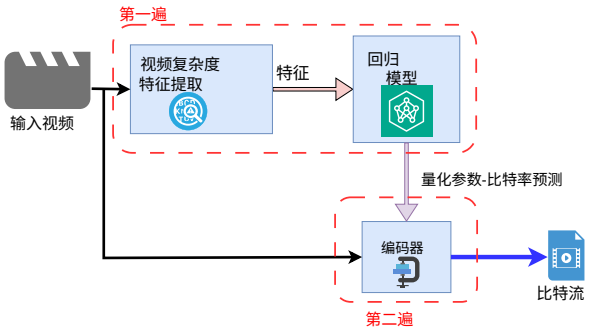


图 1：所提出的两遍码率控制框架概览，该框架将 VCA 驱动的比特预测集成到 VVenC 中。第一遍提取内容特征、预测帧级比特数，并优化第二遍编码的量化参数分配。

①提出一种机器学习框架，利用 VCA 导出特征和量化参数预测 I 帧、P 帧和 B帧的帧级比特消耗。2 验证空间复杂度和时间复杂度特征能够指导精准的比特分配，从而实现现代编码器中高能效且稳定的码率控制。

II. 所提出的两遍码率控制框架

A. 视频复杂度特征提取

任何基于机器学习的码率预测模型的有效性都在很大程度上取决于输入特征的选择。在本研究中，我们采用了七种源自视频复杂度分析器（VCA）的特征 [17]，该分析器提取轻量且具有代表性的空间和时间复杂度描述符 $\{E_{\mathrm{Y}}, L_{\mathrm{Y}}, E_{\mathrm{U}}, L_{\mathrm{U}}, E_{\mathrm{V}}, L_{\mathrm{V}}, h\}$，其中 E 量化空间纹理复杂度，L 表示亮度，而 h 则测量帧间纹理变化，将其作为时间复杂度指标 [17]。

如图 2 所示，图像组（GOP）结构遵循分层双向预测模式。为捕捉与该层级结构相一致的时间依赖关系，VCA 会以多个时间采样间隔运行。对每一帧以及采用子采样配置的帧进行特征提取，这些配置分别分析每隔 2、4、8、16 或 32 帧的内容，对应的时间间隔为 1、2、4、8、16 和 32 帧。这种多尺度提取能够捕捉不同层级上的短期和长期时间依赖关系。较小的间隔所提取的特征侧重于快速运动或场景变化，而较大的间隔则捕捉缓慢或全局变化，从而提升模型在多种时间动态下的泛化能力。

B. 回归模型

1) I 帧：I帧采用帧内编码，不参考其他帧，其码率主要取决于空间复杂度 [23]。因此，编码一个 I 帧所需的比特数（b_I）建模如下：

$$\hat{b}_I = f(E_{\mathrm{Y}}, L_{\mathrm{Y}}, E_{\mathrm{U}}, L_{\mathrm{U}}, E_{\mathrm{V}}, L_{\mathrm{V}}, q), \tag{1}$$

其中，b_I 表示为其 VCA导出特征和量化参数（q）的函数。

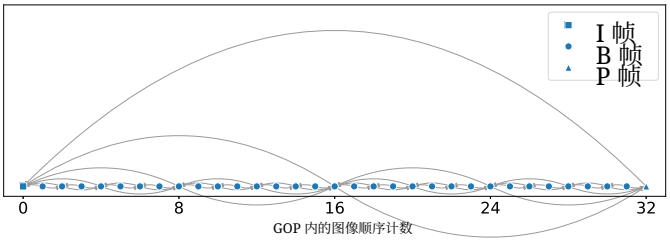


图 2：展示帧间预测依赖关系的分层图像组-32结构（0—I 帧，32—P 帧，其余为 B 帧）。

2) P 帧：P帧利用帧间预测，参考过去的帧。因此，其比特消耗不仅取决于空间特征，还取决于反映运动以及相对于参考帧的差异的时间特征。因此，P 帧的编码比特数（b_P）建模如下：

$$\hat{b}_P = f(E_{\mathrm{Y}}, h_{ref}, L_{\mathrm{Y}}, E_{\mathrm{U}}, L_{\mathrm{U}}, E_{\mathrm{V}}, L_{\mathrm{V}}, q, q_{ref}), \tag{2}$$

其中，h_{ref}量化当前帧与参考帧在亮度平面上的纹理能量差异，q_{ref} 表示参考帧的量化参数。

3) B 帧：B帧采用双向预测，同时参考过去帧和未来帧。如以下公式所示，其比特消耗取决于空间复杂度、相对于两个参考帧的时间梯度（h_{ref_1} 和 h_{ref_2}），以及当前帧和参考帧的量化参数（q_{ref_1} 和 q_{ref_2}）：

$$\begin{aligned}\hat{b}_B = f(&E_{\mathrm{Y}}, h_{ref_1}, h_{ref_2}, L_{\mathrm{Y}}, E_{\mathrm{U}}, L_{\mathrm{U}},\\ &E_{\mathrm{V}}, L_{\mathrm{V}}, q, q_{ref_1}, q_{ref_2}).\end{aligned} \tag{3}$$

尽管B帧通常比 I 帧或 P 帧消耗更少的比特，但由于存在多重依赖关系，我们预计其预测难度更高。

C. 第二遍

在第二遍编码中，利用第一遍的比特预测来优化每一帧的量化参数 (QP) 分配。第 II 节所述的第一遍码率–量化参数估计器提供初始量化参数值 q 和相应的比特预测 b。利用这一数据对以及第二遍的目标比特数 b'，VVenC 的内部 R–QP 模型 [24] 会确定调整后的量化参数 q'，以尽可能满足目标码率。

如第 [24], 节所述，初始估计值 $\bar{q}$ 的计算如下：

$$\bar{q} = q - c_{\text{low}} \cdot \sqrt{\max(1; q)} \cdot \log_2\left(\frac{b'}{\hat{b}}\right), \tag{4}$$

其中，c_{low} 是与 R–QP 曲线的低码率区域相关的缩放常数。为提高较高比特率下的稳健性，还会加入一个校正项：

$$q' = \text{round}\big(\bar{q} + c_{\text{high}} \cdot \max(0; q_{\text{start}} - \bar{q})\big), \tag{5}$$

其中，c_{high} $(0 < c_{\text{high}} < 1)$ 是一个分辨率相关常数（例如，对于 2160p 内容为 0.5，对于 480p 内容为 0.25），而 $q_{\text{start}} = 24$

是通过经验选取的参考点。

该流程确保在最终进行率失真优化编码之前，将预测的比特分配映射到最合适的量化参数值。此遍之后，实际比特数与目标比特数之间的任何剩余偏差，都会在后续帧中进行自适应补偿，如 [24] 所述。在图像组层面，目标比特率根据第一遍预测的帧级比特消耗 b 按比例分配给各帧，从而在满足整体码率约束的同时保留图像组内的相对复杂度差异。，进而为后续码率分配决策提供更可靠的依据，并有助于在同一图像组中呈现异质运动、纹理变化或突发场景转换的复杂序列中保持一致的编码行为。

III. 评估设置

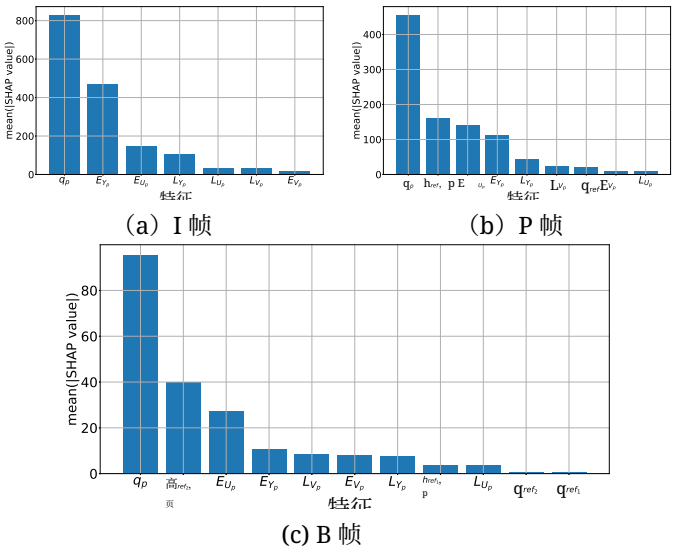
A. 第一遍

1）数据集：为训练本研究中的回归模型，我们采用 Inter-4K 数据集 [25]，，其中包含 1000 个多样化的超高清视频序列，使模型能够接触到真实场景中的特征，而不局限于 CTC 测试序列中相对稳定的场景结构。视频使用 VVenC [22] 编码，量化参数为 $\in [20, 50]$ ，采用 RA 编码结构 [26]，，图像组大小为 32 帧，帧内周期为 64 帧，并使用更快预设 [27]。为验证模型的泛化能力，我们进行了 5 折交叉验证，并对所有折的结果取平均值。

2）特征提取：使用 VCA 版本 1.5 [17] 提取视频复杂度特征。将块级离散余弦变换（DCT）特征汇总为帧级描述符。根据当前帧是 P 帧还是 B 帧，相对于一个或两个参考帧计算时间梯度特征 h。提取过程使用 8 个 CPU 线程，并采用 x86 SIMD 优化。

3）学习框架：我们采用线性回归、随机森林回归 [28]，和 XGBoost [29] 进行比特估计。我们选择这些模型，而非循环模型或基于序列的模型（例如长短期记忆网络），因为 VCA 特征已通过多尺度梯度显式编码时间依赖关系。此外，基于树的模型性能出色，同时推理延迟和内存占用显著更低。训练标签取自编码器逐帧实际使用的比特数。特征经过归一化处理，模型分别针对 I 帧、P 帧和 B帧。超参数包括估计器数量（n_estimators）、最大树深度（max_depth），以及分裂和叶节点约束。在我们的实验中，n_estimators 固定为 100，min_samples_split 和 min_samples_leaf 分别设为 2 和 1，树深度最大为 16。所有模型均使用 scikit-learn [30]实现，随机状态为=0。_

4) 性能指标： 预测精度通过以下指标进行评估：



（a）I 帧　　（b）P 帧

(c) B 帧

图 3：比特预测中特征的相对重要性。

- 平均绝对百分比误差（MAPE）：以百分比衡量平均预测误差大小。
- R^2 得分：表示模型所解释的真实值方差比例。

B. 第二遍

采用随机访问配置和更快 预设 [27]，（不使用时间滤波）评估率失真性能。虽然绝对比特消耗特性会随较慢的预设而变化，但由于所提出的内容—比特关系依赖于基础的空间—时间特征，而非特定于预设的启发式规则，因此预计该关系仍然有效。通过固定量化参数的类似CTC的编码获得预定义目标码率。使用 JVET SDR CTC序列（A1和A2类）[26]进行评估，并将其转换为 8位、30帧/秒格式。实验在运行 Linux 和 GCC 7.3.1 的 Intel Xeon E5-2697A v4 系统上进行，使用八个 CPU 线程。锚点为采用更快预设的固定量化参数随机访问 VVenC 编码。

结果采用组合 YUV Bjontegaard Delta-Rate（BD_{YUV}）、[31], 平均编码时间差（ΔT）以及与目标比特率的偏差进行报告。在视觉质量相当时，较低的 BD 码率表示更高的编码效率。

IV. 实验结果

A. 预测精度

1) 模型比较与消融研究： 表 I 总结了不同回归模型对 I 帧、P 帧和 B帧进行帧级比特预测的性能。在所评估的模型中，随机森林（RF）始终表现最佳，其 I 帧、P 帧和 B帧的 R^2 值分别为 0.93、0.88 和 0.77；因此，我们在第 IV-B 节所进行的码率控制评估中采用 RF。其在所有帧类型上的预测误差均保持在 9% 平均绝对百分比误差以下，表明其具有较强的泛化能力和稳定的学习性能。

表 I：帧级比特预测模型的性能。

模型	I 帧		P 帧		B 帧	
	MAPE [%]	R^2	MAPE [%]	R^2	MAPE [%]	R^2
线性回归 [无色度]	8.92	0.86	11.15	0.79	17.48	0.65
XGBoost [无色度]	7.81	0.89	10.04	0.83	15.72	0.70
随机森林 [无色度]	7.35	0.90	9.46	0.84	14.85	0.72
线性回归	8.51	0.87	10.72	0.80	16.82	0.67
XGBoost	7.12	0.91	8.83	0.86	13.26	0.74
随机森林	**6.84**	**0.93**	**8.21**	**0.88**	**12.94**	**0.77**

2) I 帧： 图 3a 展示了比特预测中特征的相对重要性，该重要性通过 SHAP 值 [32]进行衡量。量化参数 q 在排名中占主导地位，反映了其对帧内帧比特分配的直接控制作用。下一个影响最大的描述符是亮度纹理能量 E_{Y}，它捕捉了帧的空间频率特性。色度纹理能量（E_{U}、 E_{V}）和亮度项（L_{Y}、L_{U}、 L_{V}）的贡献较小，但对细粒度预测仍然有用。对于 I帧，随机森林模型的平均绝对百分比误差为 6.84%，且 $R^2 = 0.93$，表明其具有良好的稳健性。

3) P 帧：P 帧预测综合了空间信息与时间信息。如图 3b 所示，时间梯度特征的影响力仅次于 q ，能够捕捉帧间运动复杂度。所提出的随机森林模型的平均绝对百分比误差达到 = 8.21% 和 $R^2 = 0.88$，展现出较强的预测能力。

4）*B 帧*：B 帧预测更具挑战性，因为它采用双向参考，且运动补偿精度各异。如图 3c 所示，基于过去和未来参考帧得出的时间梯度对码率估计有显著贡献，并补充了量化参数和亮度纹理线索。B 帧模型的平均绝对百分比误差为 = 12.94% 且$R^2 = 0.77$。尽管其相关性略弱于 I 帧和 P 帧，但由于 B 帧本身消耗的比特数较少，整体误差幅度仍然很小。

5）色度特征消融：如表 I 所示，排除色度特征会使所有帧类型的性能略有下降，平均绝对百分比误差约增加 1–2%，且 R^2 降低 0.02–0.04。该影响在 B 帧中最为明显，因为色度纹理为运动补偿预测残差提供了额外信息。

B. 码率控制结果

表 II 对比了所提出的码率控制方法与传统双遍码率控制（2pRC） [24], [33]在 JVET 超高清测试序列 [26]上的表现。在随机访问模式下，所提出的方法实现了相当的编码效率，平均 BD 码率为 -0.14%，而 2pRC 为 0.26%。虽然部分序列（例如，*FoodMarket4*、*ParkRunning3*）表现出显著增益，但其他序列的 BD 码率仅略有增加。这种表现体现了预测器的内容自适应特性：它优先保证码率稳定性和码率符合性，但在内容高度动态或不规则时，偶尔会以牺牲少量编码效率为代价。对于略有编码损失的序列，其偏差仍处于码率控制机制通常的可变性范围内

表 II：在随机访问配置下，使用双遍码率控制和所提方法与固定量化参数编码的对比结果。

序列	**2pRC** BD_{YUV}[%]	**所提码率控制** BD_{YUV}[%]
Tango1	0.25	1.12
FoodMarket4	0.41	-1.83
篝火	0.18	0.26
猫机器人	0.29	-0.64
DaylightRoad2	0.36	1.47
ParkRunning3	0.10	-1.22
平均值	**0.26**	**-0.14**
ΔT [%]	**41**	**-6**
与目标码率的偏差 [%]	**0.00**	**0.05**

，并且可由编码速度和码率稳定性方面的显著提升予以弥补，这也是所提出框架的主要设计目标。总体而言，该方法无需重新编码即可实现具有竞争力的率失真性能，同时与目标比特率的偏差极小（0.05%），证实了其码率稳定性良好。

此外，所提出的框架显著缩短了编码时间。双遍码率控制（2pRC）的第一遍处理速度约为 0.40 fps，而模型驱动的处理遍次超过 10 fps，速度提升了 $25\times$倍。VCA 特征提取与随机森林推理合计占总编码时间的比例不到 5%。总体而言，与 2pRC 相比，编码时间减少了 33.33%，与单遍固定量化参数编码相比则减少了 -6%。这些速度与稳定性方面的提升，使该方法适用于实时部署。

V. 结论

本文提出了一种用于 VVC 帧级比特预测的内容驱动框架，该框架利用从 VCA 中提取的特征。通过对与编码器决策过程相匹配的空间和时间特征进行建模，所提出的基于随机森林的预测器无需试编码，即可准确估计 I、P 和 B 帧的比特消耗。在超高清视频序列上的结果表明，该方法具有很强的预测准确性（I/P/B 分别为 R^2 =0.93、0.88 和 0.77），证实轻量级 VCA 描述符能够捕获码率主导复杂度。在随机访问配置下将该方法集成到码率控制环路后，其编码效率可与传统的双遍码率控制（2pRC）相媲美：所提出的方法的平均值为 BD_{YUV} -0.14 %，而 2pRC 为 0.26%，同时保持了良好的码率稳定性（偏差为 0.05%）。在复杂度方面，学习型第一遍比 2pRC 快约 $25\times$ 倍；与 2pRC 相比，总编码时间减少了 33.3%，与固定QP锚点相比则减少了 6%。因此，所提出的方法适用于移动、嵌入式和云平台上的实时及功耗受限编码器。

未来的工作将把该框架扩展到动态GOP结构，使其能够针对内容变化进行在线模型自适应，并将其集成到VVenC的码率控制模块中，以便在生产级流媒体处理管线中实现高效编码。该框架还可支持实时编码和自适应码率流媒体，其中准确的逐帧比特预测可直接提升缓冲区稳定性和用户体验。

参考文献

[1] 思科，《思科视觉网络指数：预测与方法论，2017–2022（白皮书）》，2019。[2] T. Wiegand 等，《H.264/AVC视频编码标准概述》，*IEEE*视频技术电路与系统汇刊，第13卷，第7期，第560–576页，2003。[3] G. J. Sullivan 等，《高效视频编码（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。[4] B. Bross 等，《通用视频编码（VVC）标准概述及其应用》，*IEEE*视频技术电路与系统汇刊，第31卷，第10期，第3736–3764页，2021年10月。[5] C. R. Helmrich 等，《利用VVenC和基于预分析统计的码率封顶实现快速恒定质量视频编码》，发表于*2024 IEEE*国际图像处理会议（*ICIP*），2024年，第1823–1828页。[6] W. Fei和S.-a. Zhu，《基于柯西分布的率失真模型实现H.264的最优比特分配与码率控制》，发表于*2007 IEEE*控制与自动化国际会议，2007年，第1190–1193页。[7] Q. Wang 等，《一种用于立体视频编码的基于率失真模型的帧层码率控制算法》，电子学报，第21卷，第4期，第678–682页，2012。[8] M. Rezaei、M. M. Hannuksela和M. Gabbouj，《面向可变比特率视频的半模糊码率控制器》，*IEEE*视频技术电路与系统汇刊，第18卷，第5期，第633–645页，2008。[9] Y. Li 等，《视频编码器中用于视频质量优化的低时延窗口式码率控制方案》，发表于*2014 IEEE*声学、语音与信号处理国际会议（*ICASSP*），2014年，第7333–7337页。[10] S. Sanz-Rodrıguez 等，《一种用于超低时延高清视频编码的低复杂度并行友好码率控制算法》，发表于*2013 IEEE*国际多媒体与博览会研讨会（*ICMEW*），2013年，第1–4页。[11] J. Wang、Z. Gao和X. Zhang，《面向众核平台实时超高清视频编码器的高效并行友好码率控制》，发表于*2014 IEEE*国际多媒体与博览会研讨会（*ICMEW*），2014年，第1–6页。[12] N. Srinivasamurthy 等，《一种基于子画面的码率控制算法，用于实现嵌入式视频片上系统上的H.264高清编码器实时编码并提升视频质量》，发表于*2010 IEEE*第*4*届互联网多媒体服务架构与应用国际会议，2010年，第1–6页。[13] Y.-Y. Shen和C.-H. Kuo，《HEVC中基于视频下采样、实现一致质量的两遍码率控制》，发表于*2018 IEEE*国际多媒体与博览会（*ICME*），2018年，第1–6页。[14] A. Henkel 等，《利用子采样实现两遍视频编码中的快速第一遍处理》，发表于*2024*图像编码研讨会（*PCS*），2024年，第1–5页。[15] M. Zhou 等，《通用视频编码的最优质量控制算法》，*IEEE*广播汇刊，第68卷，第3期，第582–593页，2022。[16]国际电信联盟电信标准化部门，“P.910：多媒体应用的主观视频质量评估方法”，2021年11月。 [在线]。链接：https://www.itu.int/rec/T-REC-P.910-202111-I/en [17] V. V. Menon 等，“自适应视频流传输中的绿色视频复杂度分析与高效编码”，载于第一届绿色多媒体系统国际研讨会论文集，2023年，第16–18页。[18] V. V. Menon 等，“面向自适应直播流传输的感知JND的逐标题两遍编码方案”，《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。[19] 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8月，第942–946页，ISSN: 2076-1465。[20] V. V. Menon 等，“采用感知延迟的动态分辨率编码实现节能自适应视频流传输”，载于第三届Mile-High视频会议论文集，2024年，第21–27页。[21] A. Katsenou 等，“高效自适应VVC流传输的多目标帕累托前沿优化”，2026年。 [在线]。链接：https://arxiv.org/abs/2601 10607 [22] A. Wieckowski 等，“VVenC：一种开放且优化的VVC编码器实现”，载于2021年IEEE多媒体与博览会国际会议研讨会（ICMEW），2021年，第1–2页。[23] V. V. Menon 等，“采用低复杂度视频特征的通用视频编码全帧内码率控制”，载于2023年IEEE国际图像处理会议（ICIP），2023年，第2760–2764页。[24] C. R. Helmrich 等，“使用低复杂度XPSNR模型的视频编码视觉优化两遍码率控制”，载于2021年视觉通信与图像处理国际会议（VCIP），2021年，第1–5页。[25] A. Stergiou 和 R. Poppe，“AdaPool：用于信息保留下采样的指数自适应池化”，《IEEE图像处理汇刊》，第32卷，第251–266页，2023年。[26] F. Bossen 等，“JVET-T2010：SDR视频的VTM通用测试条件和软件参考配置”，载于第20届联合视频专家组视频编码会议，线上举行，2020年10月。[27] J. Brandenburg 等，“VVenC的帕累托优化编码配置：一种快速高效的VVC编码器”，载于2021年IEEE第23届多媒体信号处理国际研讨会（MMSP），2021年，第1–6页。[28] L. Breiman，“随机森林”，《机器学习》，第45卷，第1期，第5–32页，2001年10月。[29] T. Chen 和 C. Guestrin，“XGBoost：可扩展树提升系统”，载于第22届ACM SIGKDD知识发现与数据挖掘国际会议论文集，2016年8月，第785–794页。[30] F. Pedregosa 等，“Scikit-learn：Python中的机器学习”，《机器学习研究杂志》，第12卷，第2825–2830页，2011年。[31] HSTP-VID-WPOM，“用于评估视频编码效率实验的客观指标工作实践”，载于国际电信联盟，2020年。[32] S. M. Lundberg 和 S.-I. Lee，“解释模型预测的统一方法”，载于《神经信息处理系统进展》第30卷，2017年，第4765–4774页。 [在线]。链接：
h
t
t
p://
p
a
p
e
r
s.
nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf
[33] C. R. Helmrich 等，“适用于VVenC（一种开放式VVC编码器实现）的感知场景变化和噪声的码率控制方法”，载于2022年图像编码研讨会（PCS），2022年，第241–245页。