



研究与开发

基于广义高斯分布建模帧级变换系数的 VVC 编码失真预测模型

顾轶寅, 王鸿奎, 殷海兵

(杭州电子科技大学通信工程学院, 浙江 杭州 310018)

摘要: 通用视频编码 (versatile video coding, VVC) 采用多种高级编码工具共同实现卓越的编码性能。与高效视频编码 (high efficient video coding, HEVC) 相比, VVC 的变换系数分布 (transform coefficient distribution, TCD) 具有更尖锐的峰值。针对这一现象, 对帧级 TCD 进行概率密度函数 (probability density function, PDF) 建模, 并提出一种基于统计建模的帧级编码失真预测模型, 将帧级失真建模为 TCD 分布参数和量化参数的函数。实验结果表明, 相比于拉普拉斯分布以及柯西分布, 广义高斯分布在 TCD 概率密度拟合方面表现最佳; 基于广义高斯分布的失真预测模型的预测结果最接近实际编码失真。

关键词: VVC; 失真估计; D-Q 模型

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023088

VVC coded distortion prediction model based on frame-level transform coefficient modeling of generalized Gaussian distribution

GU Yiyin, WANG Hongkui, YIN Haibin

College of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract: In versatile video coding (VVC), a variety of advanced coding tools work together to achieve excellent coding performance. Compared with high efficient video coding (HEVC), the transform coefficient distribution (TCD) of VVC has sharper peaks. In order to solve this phenomenon, the probability density function (PDF) of frame-level TCD was modeled, and a frame-level coding distortion prediction model based on statistical modeling was proposed, which modeled frame-level distortion as a function of TCD distribution parameters and quantization parameters. The experimental results show that compared with the Laplace distribution and Cauchy distribution, the generalized Gaussian distribution has the best performance in TCD probability density fitting. The prediction results based on the generalized Gaussian distribution distortion prediction model are closest to the actual coding distortion.

Key words: VVC, distortion estimation, D-Q model

收稿日期: 2023-01-30; 修回日期: 2023-04-11

通信作者: 殷海兵, yhb@hdu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61972123, No.61931008, No.62202134); 浙江省“尖兵”“领雁”研发攻关计划项目 (No.2022C01068)

Foundation Items: The National Natural Science Foundation of China (No.61972123, No.61931008, No.62202134), “Pioneer” and “Leading Goose” Research and Development Program of Zhejiang Province (No.2022C01068)



0 引言

近年来,随着超高清(ultra-high definition, UHD)视频和虚拟现实(virtual reality, VR)视频的普及,视频数据量急剧增加,高效视频编码(high efficient video coding, HEVC)^[1]已不能满足未来的视频市场需求。新一代通用视频编码(versatile video coding, VVC)标准结合先进的编码工具进一步提高了视频压缩效率。相较于 HEVC,在保证相同视频质量的前提下,VVC 可节省近 50%的码率^[2]。视频编码优化是提高视频压缩效率的有效方法,准确的编码失真预测对编码优化有着指导性作用,对率失真优化^[3]、模式选择^[4]等核心编码算法研究起到关键性作用。

在过去的十几年中,变换系数分布(transform coefficient distribution, TCD)的统计建模已经得到了广泛应用,包括水印嵌入^[5]、码率控制^[6-7]和失真估计模型等。学术界已经提出许多基于变换系数分布的编码失真预测模型^[8-14]。在这些模型中,TCD 建模方法可被分为两大类。第一类将复杂的帧级 TCD 视为一个整体,采用连续的单参数概率模型进行拟合,如拉普拉斯分布(Laplace distribution, Lap)。在这些概率模型中,拉普拉斯分布形式简洁,并且可以推导出闭合形式码率和失真模型,因此它是应用最广泛的分布形式。然而,如文献[8]所述,拉普拉斯分布无法准确拟合 TCD 的尾部。广义高斯分布(generalized Gaussian distribution, GGD)^[12]随着形状参数的减小,将具有一个更加尖锐的峰值,可以准确地拟合 TCD,GGD 的控制参数复杂且难以估计,文献[13]将 GGD 分布参数固定并进一步推导出适用于 VVC 的码率估计模型,但视频帧内容复杂,仅将分布参数固定会导致较大的误差估计。第二类方法将帧级 TCD 视为由不同尺寸变换块 TCD 组成的混合分布,其关键思想是,每个大小块生成的 TCD 使用单独的单参数概率模型(如拉普拉斯分布)

进行拟合,帧级 TCD 视为不同大小块的概率密度函数(probability density function, PDF)加权之和。由于分布参数增加,混合模型能够有效地拟合 TCD。在文献[15-16]中,HEVC 中的编码单元(coding unit, CU)采用分层二叉树结构,TCD 可由不同大小的 CU 概率模型进行多参数混合建模。文献[17-18]进一步考虑不同划分深度 CU 对帧级失真模型的影响程度,根据 CU 划分深度将帧级 TCD 由 4 个不同比例的拉普拉斯分布的概率密度函数拟合得到。

VVC 采用了更复杂的帧内和帧间预测,在帧内预测中,帧内模式的数量从 HEVC 中使用的 35 种扩展到 67 种^[2],新的帧内预测编码工具,例如多参考线(multiple reference line, MRL)帧内预测和帧内预测子分区(intra sub-partitions, ISP)已被合并到 VVC 中。在帧间预测中,VVC 扩展了传统的合并预测,包括基于运动矢量差异(motion vector difference, MVD)的合并模式。此外,VVC 将 CU 中超过 32×32 部分的高频变换系数强制置零。这些新的编码工具使得 VVC 的残差系数更加稀疏,之前广泛使用的拉普拉斯分布无法适用于 VVC 的 TCD 建模。

由先前的工作可知,在 HEVC 中,基于拉普拉斯分布的 TCD 建模更广泛地用于研究工作,混合拉普拉斯分布等统计模型也被提出用于更准确地描述宏块级率失真模型建模。而在 VVC 中,覆盖每个大小变换的混合模型的 TCD 建模是不切实际的。相反,考虑拟合复杂性和参数估计,单一概率模型更实用。与 HEVC 相比,VVC 预测残差更稀疏,变换系数更集中于零附近,采用一种能够拟合尖锐峰值的模型是最优选择。高斯概率分布和柯西概率分布均有较好的拟合尖锐峰值的能力,因此本文分别采用广义高斯模型和柯西分布(Cauchy distribution, Cau)模型对帧级 TCD 进行建模。考虑 I 帧、B 帧和 P 帧在预测过程中的参考方式不同,因此分别对 I 帧、B 帧和 P 帧 TCD

进行单独的统计建模。由于广义高斯分布和柯西分布很难推导出闭合形式的失真表达式,因此对编码失真和帧级分布参数之间的关系进行统计分析拟合,推导出关于分布参数的失真表达式。最终,在帧级失真估计时采用基于滑动窗口的估计策略,当前帧 TCD 分布参数由前向帧分布参数加权得到。针对本文提出的失真估计模型,采用计算均方根误差的方法评估算法准确度,能够明显地体现本文失真估计算法的优越性。

1 TCD 建模

VVC 中采用了先进的编码工具,使变换系数分布集中在零附近。HEVC 和 VVC 的 TCD 比较如图 1 所示,与 HEVC 相比,VVC 的 TCD 具有更尖锐的峰值,同时,VVC 采用嵌套多类型树结构(multi-type tree, MTT)代替 HEVC 中多单元类型的概念^[19],并支持块形状灵活变动,这导致 VVC 中大小块数量大大增加。因此,覆盖所有大小变换的混合模型对 VVC 的 TCD 建模是不切实际的。相反,考虑拟合复杂度和参数估计难度,使用单参数概率模型更为实用。

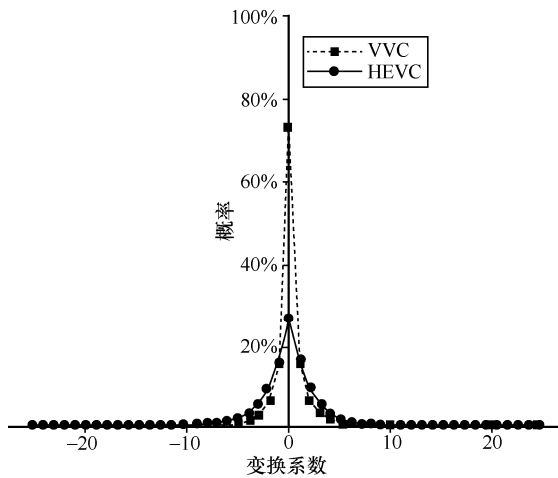


图1 HEVC 和 VVC 的 TCD 比较

VVC 中 FourePeople 序列首帧 TCD 建模如图 2 所示。从图 2 可知,广义高斯分布和柯西分布拥有更加尖锐的峰值,并且能够较好地拟合

TCD 的尾部。因此,分别采用柯西分布和广义高斯分布对具有更尖锐峰值的 TCD 进行统计建模。

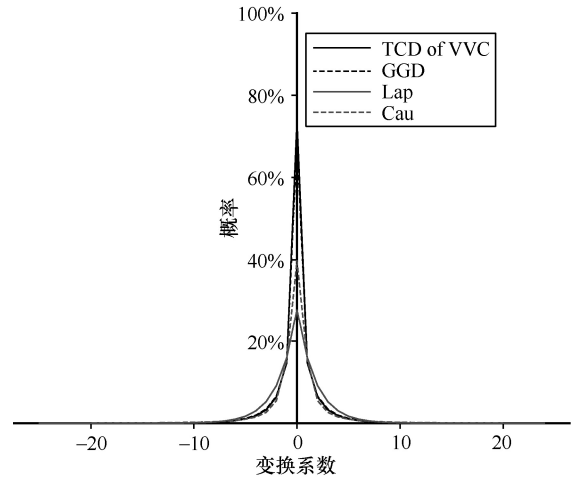


图2 VVC 中 FourePeople 序列首帧 TCD 建模

柯西分布: Cau 分布函数 $f_{\text{Cau}}(x)$ 可由式 (2) 表示。

$$f_{\text{Cau}}(x) = \frac{\beta_c}{\pi(\beta_c + x^2)}, x \in \mathbb{Z} \quad (1)$$

在大多数情况下,实际 TCD 具有相当大的尾部。具有参数 β_c 的零均值柯西分布表现出与实际交流分量(alternating component, AC)相似的行为,尾部衰减缓慢。参数 β_c 取决于视频帧内容,可以使用变换系数的统计直方图进行估计。

广义高斯分布: GGD 分布函数 $f_{\text{GGD}}(x)$ 可由式 (2) 表示:

$$f_{\text{GGD}}(x) = \frac{\alpha}{2\beta_G \cdot \Gamma(1/\alpha)} e^{-(|x|/\beta_G)^\alpha} \quad (2)$$

其中, α 为形状参数,可根据文献[14]估计。随着 α 的减小,GGD 的形状愈发尖锐。 β_G 为 GGD 的分布参数, $\beta_G = \frac{\alpha}{[2\Gamma(1/\alpha)f'(0)]}$, $f'(0)$ 是变换系数等于零的概率的导数, $\Gamma(\cdot)$ 为伽马函数,

$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ 。 α 的估计过程非常复杂,因此

本文采用 3 种具有固定形状参数的 GGD 对 I 帧、P 帧和 B 帧分别进行 TCD 建模。与 I 帧相比,P 帧、B 帧的 TCD 建模更加准确,这是因为 P 帧、B 帧



的 TCD 更加尖锐。本文根据 Mou^[13] 的统计实验结论, 将 I 帧、P 帧和 B 帧的形状参数分别设置为 0.6、0.5 和 0.35。

本文通过计算 Kullback-Leibler (KL) 离散度 D_{kl} 来评价模型的准确度, D_{kl} 由式 (3) 计算得到:

$$D_{kl} = \sum p_x \cdot \ln \frac{p_x}{\hat{p}_x} \quad (3)$$

其中, x 是变换系数的值, p_x 和 $\hat{p}_x$ 分别为实际概

率和建模概率。 D_{kl} 用于度量两个分布之间的离散度, 因此 D_{kl} 越小表明建模越精确, 此外, 采用拉普拉斯分布作为性能比较的参考模型, 不同分布模型建模准确度见表 1。从 KL 离散度实验结果来看, 广义高斯模型在 I 帧和 B 帧上的建模准确度更高, 而柯西模型在 P 帧上的建模效果更好。相较于拉普拉斯分布, 两种分布模型在 I 帧、P 帧和 B 帧上的建模准确度均有大幅提高。

表 1 不同分布模型建模准确度

序列	$D_{kl}(\text{Lap})$			$D_{kl}(\text{Cau})$			$D_{kl}(\text{GGD})$		
	I 帧	P 帧	B 帧	I 帧	P 帧	B 帧	I 帧	P 帧	B 帧
Basketball	0.796	1.043	1.132	0.281	0.158	0.274	0.008	0.114	0.136
BlowingBubbles	0.122	0.434	0.637	0.104	0.035	0.092	0.126	0.113	0.188
BQMall	0.195	1.315	1.267	0.227	0.034	0.083	0.042	0.033	0.087
PeopleOnStreet	0.381	0.412	0.447	0.087	0.013	0.057	0.038	0.040	0.056
平均值	0.376 3	0.811	0.871	0.174	0.060	0.126	0.054	0.075	0.117

2 D-Q 估计模型

在视频编码中, 失真主要来自量化部分。对于具有给定死区偏移量 δ 的量化器而言, 量化失真可根据 TCD 分布, 结合量化步长 Q_{step} 进行估计。

$$D(Q_{\text{step}}) = \sum_{k=-\infty}^{\infty} \int_{(k-\delta)}^{(k+1-\delta)} (x - k \cdot Q_{\text{step}})^2 f(x) dx \quad (4)$$

其中, k 为量化值, $f(x)$ 为 TCD 分布概率模型。从式 (4) 可知, 失真取决于分布参数 β 和量化步长 Q_{step} 。不同 Q_{step} 值的失真情况如图 3 所示。从图 3 可知, 失真与 Q_{step} 存在如式 (4) 所示的幂函数关系。

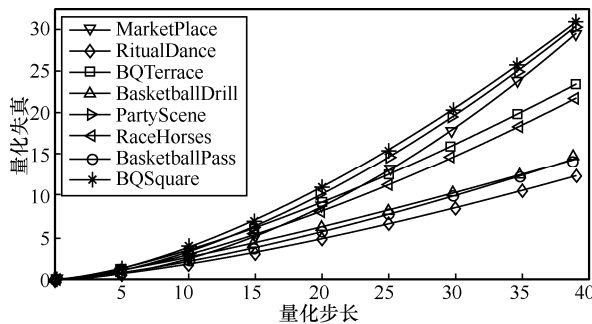


图 3 不同 Q_{step} 值的失真情况

$$D(Q_{\text{step}}) = \varphi \cdot Q_{\text{step}}^{\psi} \quad (5)$$

其中, (φ, ψ) 是失真模型参数, 每个序列对应单独的 (φ, ψ) 参数。从图 3 可以看出, 每个序列的 (φ, ψ) 代表该序列的失真特性, 而分布参数 β 也能从另一方面代表序列失真的特性。因此, 为了分析 (φ, ψ) 和分布参数 β 之间的关系, 本文将标准测试序列分成 24 组训练序列和 8 组测试序列。每一组训练序列随机采样若干帧图像, 将每一帧中 (φ, ψ) 和 β 的关系通过散点形式描绘在图 4 和图 5 中, 其中图 4 对应广义高斯分布参数 β_G , 图 5 对应柯西分布参数 β_c 。从图 4 和图 5 可以看出, φ 和 β 之间呈现指数关系, ψ 和 β 之间呈现对数关系, 因此 (φ, ψ) 可近似表示为:

$$\varphi(\beta) = \gamma_1 \cdot e^{\gamma_2 \cdot \beta} \quad (6)$$

$$\psi(\beta) = \gamma_3 \cdot \ln \beta + \gamma_4 \quad (7)$$

其中, 模型参数 $\gamma_1 \sim \gamma_4$ 的值可用表 2 中的值近似表示。

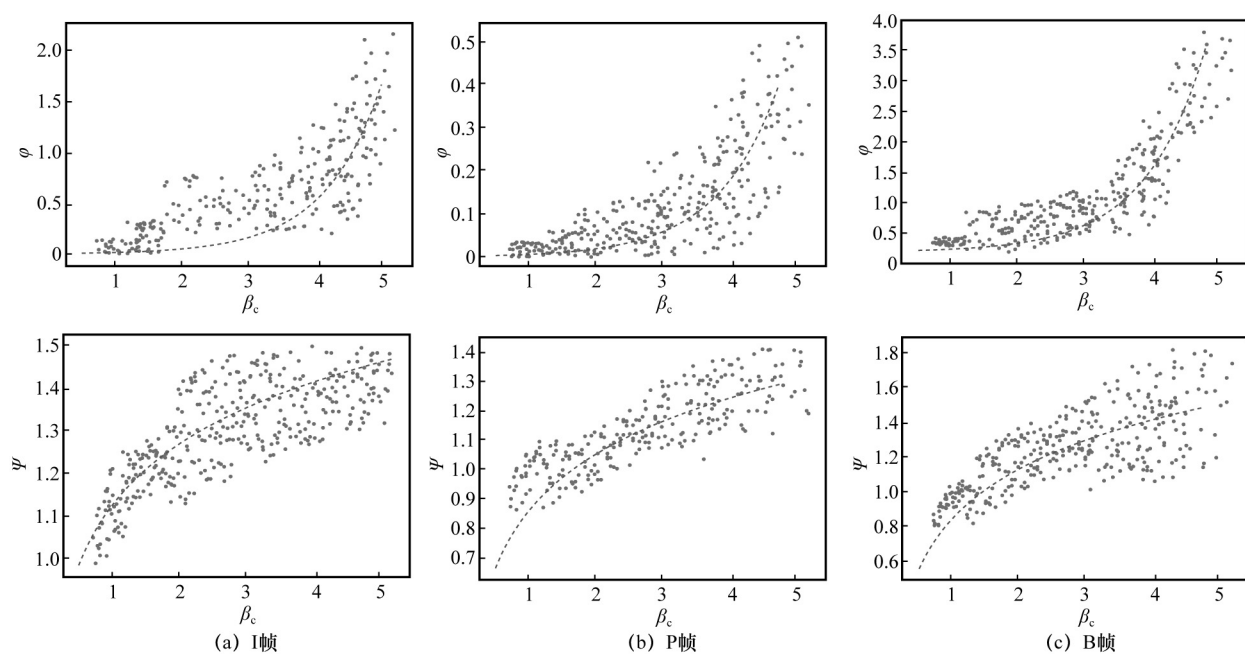
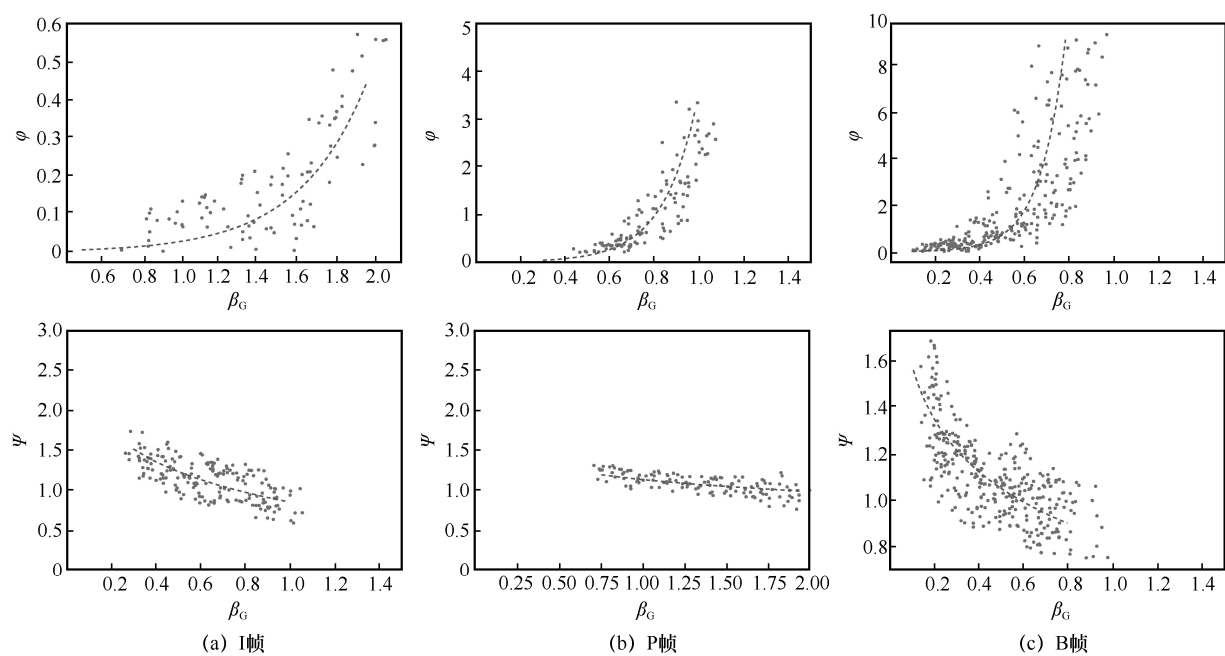
图4 模型参数 (φ, ψ) 和柯西分布参数 β_c 之间的关系图5 模型参数 (φ, ψ) 和广义高斯分布参数 β_G 之间的关系

表2 不同帧类型的模型参数

帧类型	Cau 模型				GGD 模型			
	γ_1	γ_2	γ_3	γ_4	γ_1	γ_2	γ_3	γ_4
I 帧	0.004 9	1.178 1	0.212 2	1.129 1	0.005 3	6.512 4	-0.555	0.837 4
P 帧	0.002 2	1.049 1	0.274 2	0.854 7	0.001 6	2.897 4	-0.214	1.119 7
B 帧	0.016 2	1.078 6	0.408 2	0.844 4	0.00 8	8.904 9	-0.31	0.831 7



3 基于加权平均的滑动窗口失真估计

在当前帧编码结束前, 无法获取该帧的分布参数。在实际编码过程中, 当前帧分布参数由前 N 帧的分布参数移动加权平均进行估计。因此, 第 n 帧估计分布参数由式 (8) 得到:

$$\widetilde{\beta}^n = \sum_{i=1}^N \omega_i \cdot \beta^{n-i} \quad (8)$$

其中, N 为当前帧前向参考帧的数量, ω_i 为加权重, β^{n-i} 为第 $n-i$ 帧的实际分布参数。由于 P 帧、B 帧参考已编码帧的程度不同, 因此单独对 I 帧、P 帧和 B 帧的权重系数进行有约束的回归训练。不同类型编码帧权重系数见表 3, 在编码实验中, 编码 I 帧、P 帧和 B 帧时会采用不同的权重系数。

表 3 不同类型编码帧权重系数

权重系数	I 帧	P 帧	B 帧
ω_1	0.533	0.412	0.359
ω_2	0.333	0.204	0.312
ω_3	0.144	0.384	0.329

在编码过程中为了更准确地估计帧级失真, 所提出的模型需要对实际失真值和估计失真值的差值进行补偿, 第 n 帧的最终估计失真如式 (9) 所示。

$$D_n(Q_{\text{step}}, \widetilde{\beta}^n) = \varphi(\widetilde{\beta}^n) \cdot Q_{\text{step}}^{\varphi(\widetilde{\beta}^n)} + \Delta D_n \quad (9)$$

其中, ΔD_n 为第 n 帧的失真补偿, 通过前 N 帧失真差值的加权平均得到:

$$\Delta D_n = \sum_{i=1}^N \omega_i \cdot (\hat{D}_{n-i} - \varphi(\widetilde{\beta}^{n-i}) \cdot Q_{\text{step}}^{\varphi(\widetilde{\beta}^{n-i})}) \quad (10)$$

其中, $\hat{D}_{n-i}$ 为第 $n-i$ 帧的实际编码失真值。

4 实验结果与分析

为了对本文提出的模型进行评价, 在最新的参考软件 VTM10.0 中实现了本文提出的模型。实验采用了全帧内 (all intra, AI)、随机访问 (random

access, RA) 和低延迟 P 帧 (lowdelay_P, LDP) 3 种主要配置。在 8 个测试序列中测试了所提模型的性能。如式 (3) 和式 (4) 所示, 编码失真和量化步长直接相关, 在测试实验中, 序列编码 50 帧, 采用 4 个量化参数 (quantization parameter, QP) 值 22、27、32、37 进行实验。使用均方根误差 (root mean square error, RMSE) 测量失真估计的准确度, 见表 4。

将 GGD 模型与 Cau 模型和 Lap 模型进行了比较。根据式 (4) 推导基于 Lap 模型的失真表达式为:

$$D(Q_{\text{step}}, \beta_L) = \frac{(2 + Y - 2\delta Y) Y e^{\delta Y} + 2 - 2e^Y}{(\beta_L)^2 (1 - e^Y)} \quad (11)$$

其中, $Y = \beta_L \cdot Q_{\text{step}}$, β_L 为拉普拉斯分布参数。从表 4 可以看出, 在 AI、LDP 和 RA 配置下, GGD 模型平均 RMSE 值分别为 0.046、0.050、0.166, 而 Cau 和 Lap 模型平均 RMSE 分别为 0.052、0.057、0.220 和 0.057、0.082 9、0.326。P 帧在帧间预测时仅参考前向帧进行运动估计, 本文提出模型的分布参数和失真补偿参考前向帧进行加权移动平均估计, 这使得在 I 帧和 P 帧中的估计较为准确, AI 和 LDP 配置中的 RMSE 平均值小于 RA 配置。

帧级实际编码失真和 3 种模型估计失真的实际输出峰值信噪比 (peak signal-to-noise ratio, PSNR) 比较如图 6 所示, PSNR 的计算如式 (12) 所示:

$$\text{PSNR} = 10 \cdot \lg \left(\frac{\text{MAX}^2}{\text{MSE}} \right) \quad (12)$$

其中, 对于 8 bit 表示的视频, MAX 取值为 255, MSE 表示压缩视频与原始视频之间的均方根误差。从图 6 可以看出, GGD 模型的 PSNR 估计更加准确。具体而言, GGD 模型在 AI 和 LDP 配置下具有更好的估计能力, LDP 配置下实际编码 PSNR 值会出现阶段性的波动, GGD 模型也能够较好地追踪实际 PSNR 的波动。总的来说,

表4 失真预测 RMSE 比较

序列	QP	Lap			Cau			GGD		
		AI	LDP	RA	AI	LDP	RA	AI	LDP	RA
BasketballDrive	22	0.170	0.040	0.226	0.160	0.043	0.189	0.161	0.038	0.181
	27	0.113	0.075	0.037	0.093	0.084	0.087	0.099	0.133	0.038
	32	0.117	0.103	0.157	0.124	0.111	0.143	0.107	0.092	0.243
	37	0.128	0.109	0.270	0.125	0.098	0.164	0.123	0.100	0.233
FourePeople	22	0.047	0.060	0.607	0.022	0.064	0.319	0.016	0.028	0.160
	27	0.051	0.133	1.017	0.037	0.104	0.242	0.023	0.035	0.133
	32	0.038	0.067	0.608	0.041	0.063	0.343	0.028	0.046	0.160
	37	0.041	0.058	0.318	0.044	0.042	0.435	0.038	0.049	0.190
BasketballDrill	22	0.070	0.255	0.325	0.063	0.166	0.049	0.059	0.049	0.199
	27	0.072	0.117	0.213	0.059	0.071	0.048	0.066	0.055	0.131
	32	0.068	0.074	0.155	0.069	0.064	0.065	0.065	0.053	0.123
	37	0.066	0.061	0.134	0.077	0.049	0.099	0.064	0.056	0.117
BQSquare	22	0.045	0.047	0.086	0.038	0.041	0.118	0.028	0.036	0.089
	27	0.045	0.034	0.096	0.043	0.039	0.207	0.037	0.032	0.081
	32	0.033	0.041	0.117	0.033	0.049	0.197	0.031	0.038	0.122
	37	0.037	0.055	0.184	0.046	0.042	0.097	0.037	0.054	0.198
Cactus	22	0.029	0.021	0.031	0.013	0.011	0.027	0.008	0.008	0.029
	27	0.021	0.016	0.128	0.023	0.019	0.106	0.009	0.012	0.055
	32	0.021	0.029	0.034	0.022	0.010	0.049	0.016	0.023	0.074
	37	0.027	0.035	0.106	0.026	0.029	0.084	0.025	0.034	0.099
KristenAndSara	22	0.067	0.078	0.828	0.026	0.071	0.753	0.015	0.016	0.138
	27	0.045	0.101	1.192	0.033	0.105	0.913	0.020	0.024	0.124
	32	0.037	0.077	0.756	0.064	0.062	0.642	0.026	0.029	0.115
	37	0.034	0.050	0.401	0.023	0.033	0.315	0.028	0.029	0.146
PartyScene	22	0.043	0.317	0.587	0.049	0.029	0.346	0.019	0.054	0.454
	27	0.035	0.147	0.365	0.036	0.044	0.198	0.028	0.072	0.297
	32	0.035	0.087	0.246	0.033	0.049	0.169	0.033	0.066	0.217
	37	0.038	0.058	0.182	0.041	0.068	0.044	0.038	0.050	0.177
BlowingBubbles	22	0.051	0.054	0.370	0.039	0.072	0.243	0.032	0.043	0.367
	27	0.060	0.073	0.266	0.043	0.046	0.156	0.052	0.070	0.257
	32	0.073	0.091	0.220	0.063	0.011	0.098	0.071	0.090	0.204
	37	0.077	0.092	0.188	0.064	0.049	0.101	0.076	0.092	0.186
平均值		0.057	0.082 9	0.326	0.052	0.057	0.220	0.046	0.050	0.166

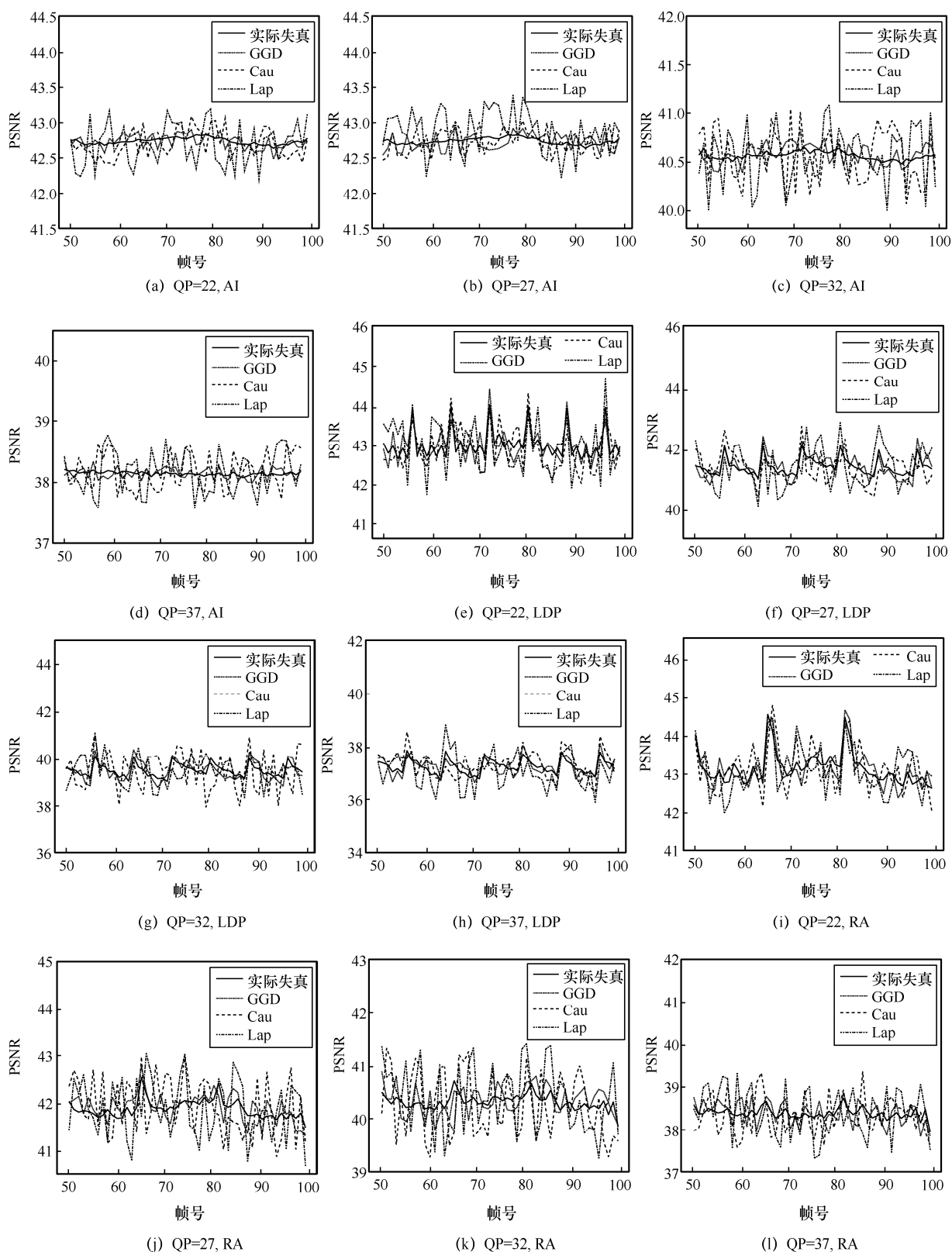


图6 帧级实际编码失真和3种模型估计失真的PSNR比较

从图 6 可以看出, 基于 GGD 模型得到的 PSNR 值更加接近实际编码 PSNR 值, 在实际编码 PSNR 波动较大的情况下, 预测值也能更加接近实际 PSNR, 说明了 GGD 模型的有效性。

在本文中, I 帧、P 帧和 B 帧的形状参数被设置为固定值, 以简化基于 GGD 的 TCD 建模过程。此外, 在分析模型参数和分布参数关系的过程中没有考虑量化效应, 这导致在实际数据和图 4、图 5 中的拟合曲线之间可以观察到建模误差, 并且误差随着 QP 的增加逐渐增大。在图 6 中同时能观察到, 随着 QP 增大, PSNR 的波动更大, 误差更加明显。因此, 在未来的工作中, 在确定模型参数时将适当考虑量化效应, 并将深入研究式 (8) 和式 (10) 中的加权值, 以便更准确地估计失真。

5 结束语

编码失真预测算法的准确度问题是率失真优化等编码模块研究的关键, 本文针对 VVC 视频编码器提出了一种基于统计建模的帧级失真预测模型。根据变换系数的分布特点, 实验分别采用 Cau 模型和 GGD 模型进行 TCD 拟合。并在此基础上采用基于滑动窗口的失真预测策略, 推导出基于帧级分布参数的统计建模失真预测模型。实验结果表明, 基于 GGD 的失真预测模型优于 Lap 模型和 Cau 模型。因此, 在未来的工作中, 将考虑基于 GGD 进一步推导出帧级码率估计模型, 结合本文失真估计模型建立准确的帧级率失真预测模型, 并对帧级 QP 进行帧级编码前的调整, 以获得更优的编码性能。

参考文献:

- [1] SULLIVAN G J, OHM J R, HAN W J, et al. Overview of the high efficiency video coding (HEVC) standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1649-1668.
- [2] BROSS B. Versatile video coding (draft 1): JVET-J1001[S]. 2018.
- [3] Qualcomm Inc. Rate distortion optimized quantization[EB]. 2008.
- [4] CHEN J, YE Y, KIM S. Algorithm description for versatile video coding and test model 10 (VTM 10.0)[EB]. 2020.
- [5] CHEN L N, YANG G B, HO A T. A Cauchy distribution based video watermark detection for H.264/AVC in DCT domain[C]//Proceedings of 2011 IEEE International Symposium of Circuits and Systems (ISCAS). Piscataway: IEEE Press, 2011: 2665-2668.
- [6] PANG C, AU O C, DAI J J, et al. LMM-based frame-level rate control for H.264/AVC high-definition video coding[J]. Signal Processing: Image Communication, 2012, 27(7): 737-748.
- [7] KANG J W. Novel distribution model of transformed coefficients in video coding using quad-tree structured block partitioning[J]. Multidimensional Systems and Signal Processing, 2017, 28(4): 1589-1609.
- [8] SI J J, MA S W, WANG S Q, et al. Laplace distribution based CTU level rate control for HEVC[C]//Proceedings of 2013 Visual Communications and Image Processing (VCIP). Piscataway: IEEE Press, 2014: 1-6.
- [9] HANG H M, CHEN J J. Source model for transform video coder and its application. I. Fundamental theory[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(2): 287-298.
- [10] ALTUNBASAK Y, KAMACI N. An analysis of the DCT coefficient distribution with the H.264 video coder[C]//Proceedings of 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing. Piscataway: IEEE Press, 2004.
- [11] KAMACI N, ALTUNBASAK Y, MERSERAU R M. Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(8): 994-1006.
- [12] ZHAO X, SUN J, MA S W, et al. Novel statistical modeling, analysis and implementation of rate-distortion estimation for H.264/AVC coders[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(5): 647-660.
- [13] LI Y, MOU X Q. Transform coefficients distribution of the future versatile video coding (VVC) standard[C]//Proceedings of SPIE/COS Photonics Asia, SPIE 10817, Optoelectronic Imaging and Multimedia Technology V. [S.l.:s.n.], 2018: 120-130.
- [14] SUN J, DUAN Y Z, LI J T, et al. Rate-distortion analysis of dead-zone plus uniform threshold scalar quantization and its application: part I: fundamental theory[J]. IEEE Transactions on



- Image Processing: a Publication of the IEEE Signal Processing Society, 2013, 22(1): 202-214.
- [15] YANG E H, YU X, MENG J, et al. Transparent composite model for DCT coefficients: design and analysis[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2014, 23(3): 1303-1316.
- [16] GAO W, KWONG S, YUAN H, et al. DCT coefficient distribution modeling and quality dependency analysis based frame-level bit allocation for HEVC[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(1): 139-153.
- [17] LEE B, KIM M. Modeling rates and distortions based on a mixture of Laplacian distributions for inter-predicted residues in quadtree coding of HEVC[J]. IEEE Signal Processing Letters, 2011, 18(10): 571-574.
- [18] CUI J, WANG S S, WANG S Q, et al. Hybrid Laplace distribution-based low complexity rate-distortion optimized quantization[J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2017, 26(8): 3802-3816.
- [19] BROSS B, CHEN J, SHAN L. Versatile video coding (draft 2): JVET-K1001[S]. 2018.

[作者简介]



顾轶寅(1997-), 男, 杭州电子科技大学通信工程学院硕士生, 主要研究方向为视频编解码。



王鸿奎(1990-), 男, 博士, 杭州电子科技大学通信工程学院讲师, 主要研究方向为感知视频编码。



殷海兵(1974-), 男, 博士, 杭州电子科技大学通信工程学院教授, 主要研究方向为数字视频编解码。